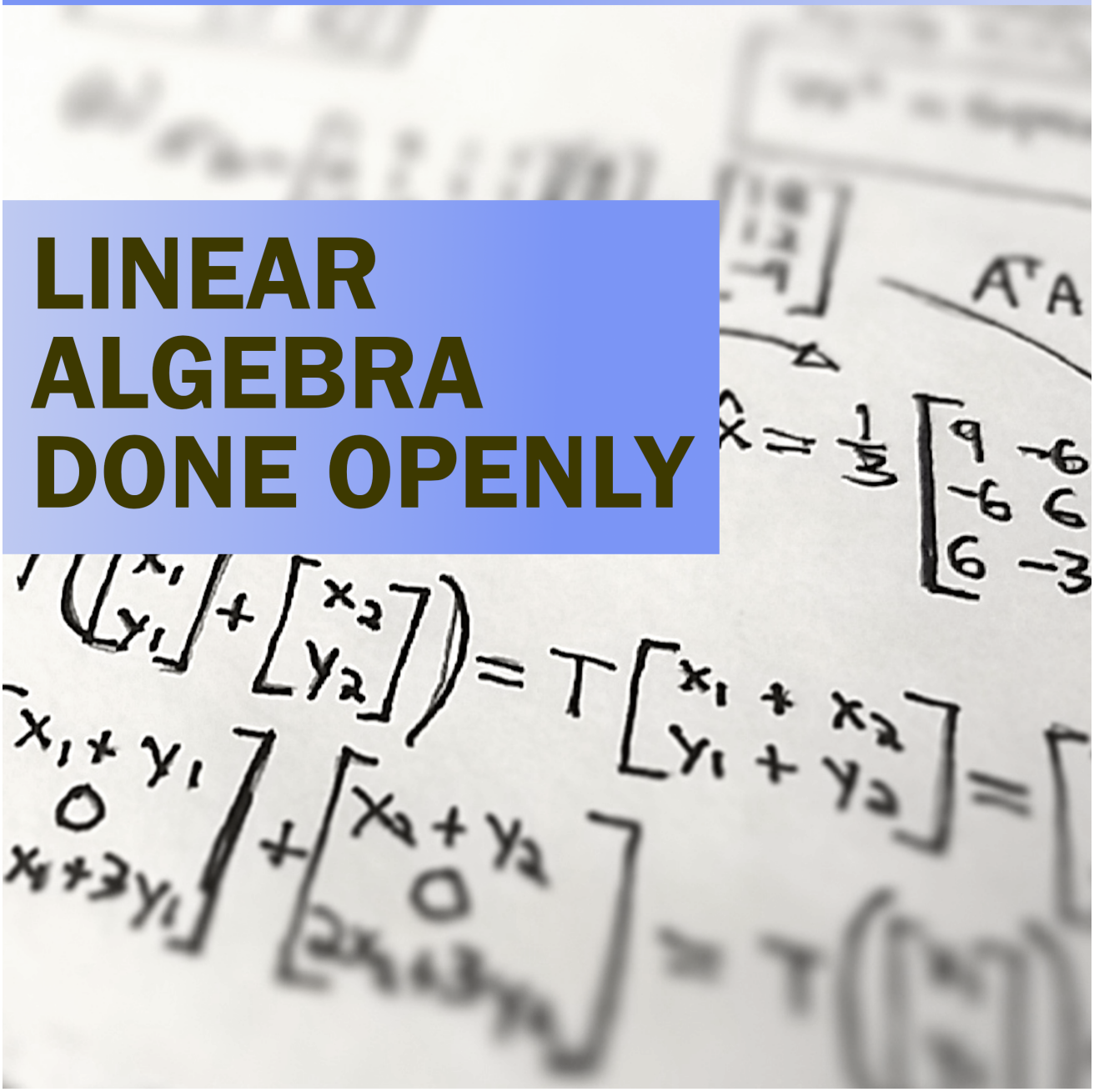
Andrew Misseldine

LINEAR ALGEBRA DONE OPENLY



Linear Algebra Done Openly

Andrew Misseldine

June 30, 2023

Contents

Preface **5**

1 Introduction to Linear Algebra **1**
1.1 Linear Systems . . . 2
1.2 Fields . . . 8
1.2.1 Fields . . . 8
1.2.2 Modular Arithmetic and Finite Fields . . . 9
1.3 Vector Spaces . . . 14
1.4 Linear Transformations . . . 22
1.5 Augmented Matrices . . . 28
1.6 Reduction of Linear Systems . . . 37

2 Vectors **45**
2.1 Vector Equations . . . 46
2.2 Matrix Equations . . . 53
2.3 Linear Independence . . . 59
2.4 Affine Geometry . . . 65
2.5 Subspaces . . . 71
2.6 Solution Sets of Linear Systems . . . 75
2.7 Bases . . . 81
2.8 Coordinates . . . 88

3 The Algebra and Geometry of Matrices **95**
3.1 Matrix Operations . . . 96
3.2 Matrix Properties . . . 102
3.3 Matrix Inverses . . . 106
3.4 Elementary Matrices . . . 112
3.5 Matrix Factorizations . . . 116
3.5.1 Generalizations of Elementary Matrices . . . 116
3.5.2 The LU Factorization . . . 119
3.6 Linear Transformations on $\mathbb{R}^2$ . . . 124
3.7 Representations of Linear Transformations as Matrices . . . 131

4 Orthogonality **135**
4.1 Inner Products . . . 136
4.2 Orthogonality . . . 141
4.3 Outer Products . . . 146
4.4 Affine Transformations . . . 152
4.5 Orthogonal Projections . . . 159
4.6 The Fundamental Theorem of Linear Algebra . . . 164
4.7 The Gram-Schmidt Algorithm . . . 173
4.8 The Least Squares Problem . . . 177

5 Determinants **183**
5.1 Introduction to Determinants 184
5.2 Properties of Determinants 189
5.3 Cramer's Rule 195
5.4 Cross Products 199

6 Eigenvalues **205**
6.1 Eigenvalues and Eigenvectors 206
6.2 The Characteristic Polynomial 211
6.3 Diagonalization 215
6.4 Orthogonal Diagonalization 219
6.5 Similarity and Linear Transformations 224

Appendix A Notation Index **229**

Appendix B The Substitution and Elimination Methods **231**

Appendix C Proofs of Theorems **235**

Appendix D Solutions to Exercises **239**

Preface

List of Contributors

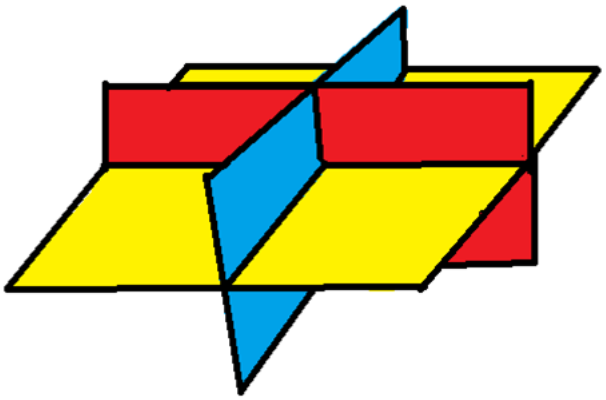
Kory Adams
Abby Allen
Will Allen
Kaden Allred
Samuel Andersen
Caroline Ashton
Alicyn Astle
Sherie Ayag
Shelby Bartlett
Tyler Bayn
Landry Benimana
Andrew Biskey
Carson Blickenstaff
Alexis Borell
Braden Carlson
Hailey Checketts
Skyler Clark
Mariah Clayson
David de Lera Santo
Charles Del Quandro
Cameron Dix
Kaylee Dockter
Riley Drishinski
Joshua Edgel
Kaden Empey
Candance Fehr
Courtney Flanigan
Jaren Frandsen
C.J. Giacoletto
Phillip Goins
Jaimie Goldberg
Caleb Goodrich
Jordan Griffith
Kaylee Hall
Malcolm Hanks
Thayne Hansen
Yingjie He
Bridger Hildreth
Christopher Houston
Ethan Hoyer
Jacob Jensen
Sofia Jones
Cyrus Kaveh
Jacob Kuhn
Maria Langford
Louaivasa Laulu
Laura Lee
Yucheng Long
Jaxton Maez
Adam Maxwell
Ellie McReaken
Kalvin Mudrow
Jacob Newey
Christopher Newton
Anthony Nguyen
Yinglong Niu
Gordon Ochsner
Brittany Palmer
Seth Palmer
Darby Parise
Colin Reid
Zach Rogers
Rindy Roos
Hamza Samha
Lucas Shaner
Hannah Simonson
Joseph Spendlove
Noah Swenson
Ashley Taylor
Jaden Torgerson
Daven Triplett
Runtian Tu
Gage VanDyke
Allyson Vest
Grayson Walker
Gregory Walsh
Sarah Walters
Adym Warhurst
Leon Weingartner
Kordell Welch
Sarah Wilcox
Walt Williams
Chance Witt
Kyle Wood
Kennedy Worthington
Jiazheng Yan
Jianhe Yu
Mattie Zeigler
Yifan Zhu
Heming Zu
Mitchell Zufelt

Chapter 1

Introduction to Linear Algebra

In this first chapter, we want to introduce all the major players in the linear arena. Linear Algebra is the study of *linear* objects (a term we are not quite ready to define). As we have seen linear objects already many times in algebra and calculus, Linear Algebra will take a detailed look at linear structures and this first chapter will explore the important fundamentals of the linear universe.

Please be aware that this first chapter may feel awkward or difficult when you read it for the first time. This is both expected and intentional. If you were familiar with all the ideas behind Linear Algebra then you probably would not be reading this book. The struggle herein will produce questions, all of which we will answer in later chapters. Consider this chapter your baptism by fire into all things linear.



"Education is that whole system of human training within and without the school house walls, which molds and develops men." – W. E. B. Du Bois

1.1 Linear Systems

In this section, we introduce the study of linear systems motivated by geometry. Thus, all numbers in this section will be real numbers. We introduce the notion of a system of linear equations and the nature of their solution sets. We discuss when a linear system is consistent/inconsistent, independent/dependent, underdetermined/overdetermined, and homogeneous/nonhomogeneous. We discuss the difference between free and dependent variables in a linear system.

Definition 1.1.1. A **system of linear equations** (or a **linear system**) is a set of equations with common variables of the following form:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b,$$

where b and the **coefficients** $a_1, \ldots, a_n$ are fixed numbers.

A **solution** to a system is an assignment to each variable such that each equation is satisfied. The **solution set** is the set of all solutions to a system of equations. Two systems of equations are **equivalent** if they have the same solution set.

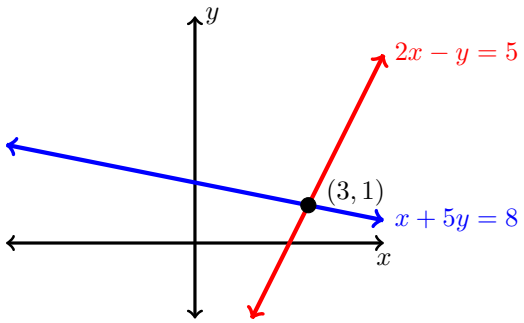
A linear equation of two real variables (typically x, y or x_1, x_2) geometrically forms a *line* in the plane. Likewise, a linear equation of three real variables (typically x, y, z or x_1, x_2, x_3) geometrically forms a *plane* in 3-space. Although much harder to visualize, a linear equation in four variables (typically x, y, z, w or x, y, z, t or x_1, x_2, x_3, x_4) geometrically forms a *hyperplane* in 4-space. Higher dimensional analogues likewise exist. As such, emphasis in this section will be placed on linear systems with two or three real variables.

Example 1.1.2. Test that $(3, 1)$ is a solution to the system

$$\begin{cases} x + 5y = 8 \\ 2x - y = 5. \end{cases}$$

Notice that $3 + 5(1) = 3 + 5 = 8$ and $2(3) - 1 = 6 - 1 = 5$. Thus, $(3, 1)$ is a solution to the above system.

Geometrically, the solution to the previous system is an intersection between two lines, as depicted to the right.



■

Example 1.1.3. Consider the linear system with three variables and three equations:

$$\begin{cases} x + 2y + 3z = 20 \\ -2x + y = -1 \\ -3x - 6y + 5z = -4. \end{cases}$$

We can check that $(2, 3, 4)$ is a solution of this system. Note that $(2) + 2(3) + 3(4) = 2 + 8 + 12 = 20$ ✓, $-2(2) + (3) = -4 + 3 = -1$ ✓, and $-3(2) - 6(3) + 5(4) = -6 - 18 + 20 = -4$ ✓, which shows that $(2, 3, 4)$ is in fact a solution of the system. Try as one might, another solution to this system CANNOT be found, that is, $(2, 3, 4)$ is the unique solution to this system of equations. ■

Example 1.1.4. Consider the linear system with three variables and three equations:

$$\begin{cases} 2x - y + z = 11 \\ x + 3y - 10z = 2 \\ -3x + 2y - 3z = -17. \end{cases}$$

It is simple to check that $(5, -1, 0)$ is a solution to this system. Likewise, we can also show that $(6, 2, 1)$ is a solution to this same system. Thus, it is possible to have multiple solutions to a linear solutions. ∎

Because of Example 1.1.4, it is worth investigating how many solutions a linear system may have. The beginning of our investigation is the following definitions.

Definition 1.1.5. A system of equations is **consistent** if it has a solution. Otherwise, we call the system **inconsistent**.

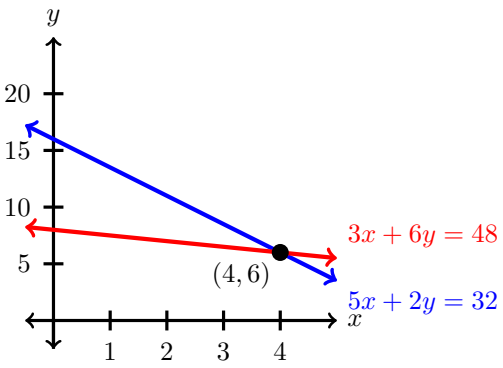
If a consistent linear system has a unique solution, we call this the **independent case**. If a consistent linear system has multiple solutions, we call this the **dependent case**.

Like we did in Example 1.1.2, we will investigate linear systems with two variables and their corresponding geometry.

Example 1.1.6. Solve each system of equations.

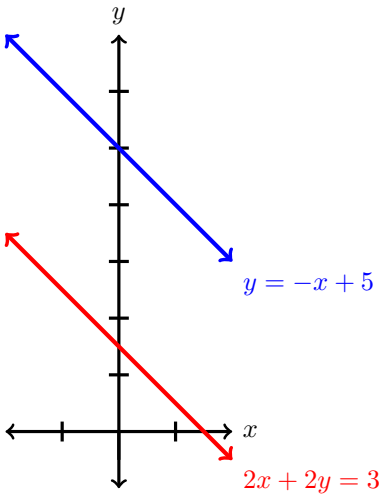
(a) $\begin{cases} 5x + 2y &= 32 \\ 3x + 6y &= 48 \end{cases}$

After graphing the two lines, as depicted to the right, we see that the two lines intersect at a unique point. This point of intersection, $\boxed{(4, 6)}$, means the system has a unique solution. Therefore, the system is consistent and independent.



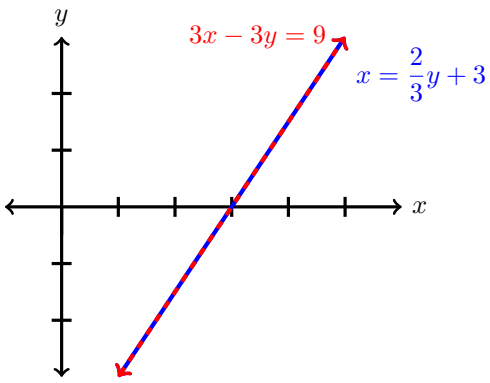
(b) $\begin{cases} y &= -x + 5 \\ 2x + 2y &= 3 \end{cases}$

After graphing the two lines, as depicted to the right, we see that the two lines are parallel. Thus, they do NOT ever intersect. Therefore, the system is $\boxed{\text{inconsisitent}}$, that is, there is no solution to the linear system.



(c) $\begin{cases} x = \frac{2}{3}y + 3 \\ 3x - 3y = 9 \end{cases}$

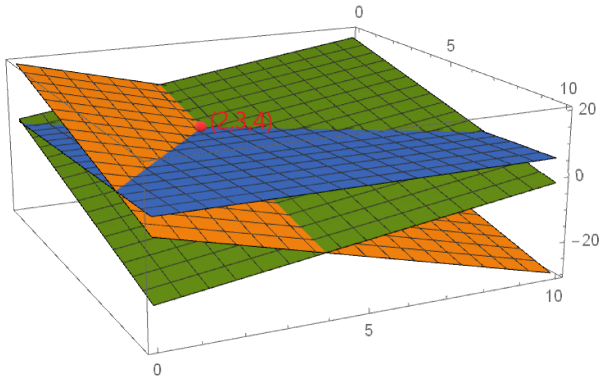
After graphing the two lines, as depicted to the right, we see that the two lines are the same line. Thus, every point on one line is also a point on the other line. Therefore, the system is dependent, that is, there are infinitely many solutions to the linear system. In particular, the system is consistent.



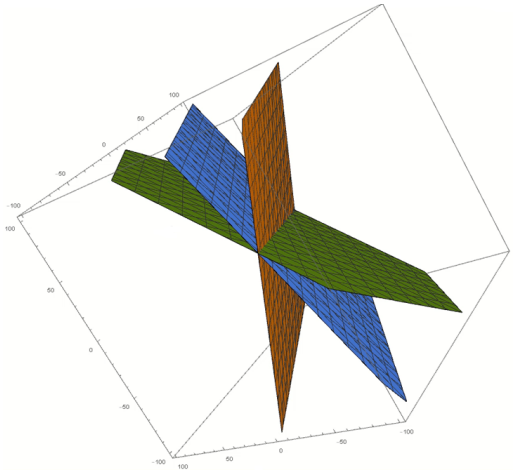
■

In general, any two lines in the real plane intersect at 0, 1, or infinitely many points (this last case occurs if the lines overlap). This holds also for higher dimensional systems of linear equations, that is, for any system of linear real equations the number of solutions is 0, 1, or ∞ and the solution set of a linear system will fall under one of these three types.

These three possibilities are also the case for intersections of planes in 3-dimensional space. The *independent, consistent* case occurs when all planes in the system intersect at a unique point. This was the case for Example 1.1.3. Visually, this can be seen as three planes intersecting at a unique point, such as the three planes involved in Example 1.1.3, illustrated below:



The *inconsistent* case occurs when not all the planes simultaneously intersect at a common point, perhaps because some subset of planes are parallel with one another. Finally, the *dependent, consistent* case occurs when the collection of planes overlay at more than one point, e.g. three planes in 3-space intersect along a common line, as illustrated below:



This was the case for Example 1.1.4. In fact, any point of the form

$$(z + 5, 3z - 1, z), \tag{1.1.7}$$

where the variable z can be freely assigned to any real number, is a solution to this linear system. This is called the **general solution** to the system. In this example, z is a **free variable**, that is, a variable which can be assigned any value. The other two variables, x and y, are called **dependent variables**, because their assignment is dependent upon the assignments of the free variables corresponding to some linear relationship. Thus, the linear system in Example 1.1.4 has one free variable and two dependent variables. The two solutions listed in Example 1.1.4 arose exactly from setting $z = 0$ and $z = 1$, respectively. In general, the dependent case occurs exactly when the linear system has at least one free variable.

The following properties will be proven in the homework.

Proposition 1.1.8. *Suppose that a consistent linear system has m equations and n variables.*

(i) If $n > m$, then the system has at least $n - m$ free variables which implies that it has multiple solutions.

(ii) If $m > n$, then the system has at least $m - n$ many equations which could be removed without modifying the solution set.

The situation where $n > m$ is called an **underdetermined system** because there are not enough equations to guarantee a unique solution. The situation where $m > n$ is called an **overdetermined system** because there might be too many equations to guarantee consistency.

Definition 1.1.9. A system of linear equations is called **homogeneous** if each equation is of the form:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = 0.$$

Otherwise, the system is called **nonhomogeneous**.

Example 1.1.10. The below 3×3 linear system is homogeneous since the right hand side of each equation is 0:

$$\begin{cases} x + 3y + 5z = 0 \\ 2x + 4y + 6z = 0 \\ 4x + 2y = 0 \end{cases}$$

On the other hand, the below linear system appears to be homogeneous at first, but is really non-homogeneous:

$$\begin{cases} x + 3y - 2z + 4 = 0 \\ 4x + 2y + 3z - 5 = 0 \\ -x + 5y - 3z - 2 = 0 \end{cases}$$

The issue here is that the linear system is not in *standard form*, that is, all the variables are located on the left-hand sides of the equations, in descending order, and the constants are all on the right-hand sides of the equations. In standard form, the system would appear as:

$$\begin{cases} x + 3y - 2z = -4 \\ 4x + 2y + 3z = 5 \\ -x + 5y - 3z = 20 \end{cases}$$

In standard form, it is much more clear that the above linear system is non-homogeneous. ∎

Proposition 1.1.11. *A homogeneous system of equations is always consistent.*

Exercises

(Go to Solutions)

For Exercises 1–4, determine if the point is a solution to the following linear system

$$\begin{cases} 2x - 3y + z = 0 \\ y + z = 0 \\ -x - 2z = 0. \end{cases}$$

1. $(0,0,0)$ 2. $(4,2,-2)$ 3. $(3,2,0)$ 4. $(1,1,1)$

For Exercises 5–8, determine if the point is a solution to the following linear system

$$\begin{cases} x + y = 7 \\ 3x + 3y - 10z = -79 \\ x + 5y + z = 33. \end{cases}$$

5. $(4,3,10)$ 6. $(0,0,0)$ 7. $(3,4,10)$ 8. $(-10,3,-38)$

For Exercises 9–12, determine if the point is a solution to the following linear system

$$\begin{cases} x + 3y - 2z = 13 \\ 2x - y = 2 \\ -x + 6y - 4z = 17. \end{cases}$$

♠ 9. $(6,10,4)$ ♠ 10. $(5,8,8)$ ♠ 11. $(3,4,1)$ ♠ 12. $(9,3,2)$

For Exercises 13–17, determine if the point is a solution to the following linear system

$$\begin{cases} x - y + z = 5 \\ 3x + 4y - 5z = -2 \\ -2x - 4y + 7z = 11. \end{cases}$$

13. $(1,2,3)$ 14. $(3,2,1)$ 15. $(5,0,6)$ 16. $(3,1,3)$ 17. $(9,3,2)$

For Exercises 18–27, graph the linear system and determine if it is consistent or inconsistent. If consistent, determine which whether it has a unique solution or multiple solutions.

18. $\begin{cases} 14x + 3y = 4 \\ 5x + 9y = 1 \end{cases}$

19. $\begin{cases} 2x - 5y = 10 \\ \frac{4}{5}x - 2y = 4 \end{cases}$

20. $\begin{cases} x = 5y - 4 \\ \frac{1}{5}x = y + 13 \end{cases}$

21. $\begin{cases} 2x + 3y = 4 \\ x - y = 2 \end{cases}$

22. $\begin{cases} -4x + y = 7 \\ -8x + 2y = 2 \end{cases}$

23. $\begin{cases} x - y = 1 \\ 2y = 2x - 2 \end{cases}$

24. $\begin{cases} y = 2x + 7 \\ y = -4x - 2 \end{cases}$

25. $\begin{cases} y = 5x + 8 \\ -5x + y - 2 = 6 \end{cases}$

26. $\begin{cases} y = 3x - 7 \\ y - 3x = 12 \end{cases}$

27. $\begin{cases} 9x + 5 = 2y \\ x + 2y = 5 \\ y = 2.5 \end{cases}$

For Exercises 28–37, determine if the linear system is homogeneous. For those which are, find a solution.

28. $\begin{cases} 4x + 2y = 0 \\ 5x + 3y = 0 \end{cases}$

29. $\begin{cases} 6x - 7y = 3 \\ 2x + 8y = 0 \end{cases}$

♠ 30. $\begin{cases} 3x - y = 0 \\ 2x + 5y = 0 \end{cases}$

♠ 31. $\begin{cases} 2x - 4y = 0 \\ -6x + 8y = 3 \end{cases}$

32. $\begin{cases} -5 - y - 2 = 0 \\ x + y + 5 = 0 \end{cases}$

33. $\begin{cases} 2x + y = 10 \\ -7x - 3y = 3 \end{cases}$

34. $\begin{cases} 5x + 2y + 3z = 2 \\ 3x - 4z = -1 \\ 5y + 3z = 5 \end{cases}$

35. $\begin{cases} 5x + 2y + 3z = 2 \\ -3x - 2y + z = 5 \\ 2x + 2y - 4z = 1 \end{cases}$

♠ 36. $\begin{cases} x - 2z = 0 \\ 3x + 4y + 4z + w = 0 \\ -2x + 6y + 6z + 2w = 0 \end{cases}$

37. $\begin{cases} 2x - 2y - 2z = 6 \\ -x - 3y + 9z = 1 \\ 4x + 3y - 18z = 3 \end{cases}$

38. Is the linear system $\begin{cases} x + 2y - 4z = 0 \\ -y + 3z = 0 \\ 2x + 3y = 0 \end{cases}$ consistent? Why or why not? Draw a 3-dimensional sketch of these planes and their intersection.

♠ 39. Using the general form for Example 1.1.4 provided on page 4 (namely (1.1.7)), construct five other solutions to the linear system from Example 1.1.4. Answers may vary.

♠ 40. Prove Proposition 1.1.11.

"Hold fast to dreams... For when dreams go
Life is a barren field... Frozen with snow." – Langston Hughes

1.2 Fields

The topic of this section will be about what is even a number. This discussion is often postponed to more advanced mathematics courses where abstraction is king, but for linear algebra students, this abstraction is just budding. This course will be primarily a computational introduction to linear algebra with select proofs and applications. In spite of this goal, the abstract notion of a field will not hinder this. In fact, the introduction of various fields will strengthen the student's understanding of the concepts and computations of linear algebra as the student will see what parts are truly necessary and when variability is allowed. This will also strengthen the realization that discussion of complex vector spaces and real vector spaces are not two separate discussion but instead one same story.

What really is a number? Although numbers are useful for counting, they are much more useful than that. A simple answer would be to say it is a thing with which we do math on. Although a crude response, this will satisfy our inquiry, thus avoiding a long journey through deep logic and philosophy. Maybe another day.

A **set** to us will mean a collection of objects, typically called **elements**. Those elements could be numbers, colors, people, pokémon, or other sets! Typically, sets will be denoted by capital alphabet letters such as A, B, C, etc. and elements with lower case letters such as a, b, c, etc. Typically, $\{\ldots\}$ is used to denote the elements in a set, i.e. $A = \{1, 2, 3, 4\}$. Often a set is described by some rule, such as $A = \{x \mid \text{x satisfies some rule}\}$. If an element x is a member of the set X, then we denote this as $x \in A$. Then notation $x \notin A$ means that x is not an element of A. There must be a clear rule to decide whether an element is a member of a set or not, otherwise it is not a set. The set which contains no elements, $\emptyset = \{\}$, is called the **empty set**.

A **function** f is a relationship between sets, say A and B, such that each element of A is assigned to exactly one element of B (although not all elements of B must be assigned to an element of A). We denote this function relation as $f : A \to B$. If A and B are two sets, we let $A \times B$ denote the set of ordered pairs of elements from A and B. For example, if $a \in A$ and $b \in B$, then $(a, b) \in A \times B$. Finally, an **operation** is a function of the form $f : A \times B \to C$. One should think of an operation as a process of bringing two objects together and creating a third object.

1.2.1 Fields

We are now ready for the definition of a field of numbers. Do not panic. If this seems very fast, that is okay. Our discussion will slow down quickly.

Definition 1.2.1. A **field** is a nonempty set F with elements called **scalars** on which are defined two operations, called **addition** $+ : F \times F \to F$ and **multiplication** $\cdot : F \times F \to F$, such that for all scalars $a, b, c \in F$, the following ten axioms hold:

(i) $a + b = b + a$

(ii) $(a + b) + c = a + (b + c)$

(iii) There exists a scalar $0 \in F$, such that $a + 0 = 0 + a = a$

(iv) For each a, there exists a scalar $-a \in F$, such that $a + (-a) = (-a) + a = 0$

(v) $a \cdot (b + c) = (ab) + (ac)$

(vi) $ab = ba$

(vii) $(ab)c = a(bc)$

(viii) There exists a scalar $1 \in F$, such that $a1 = 1a = a$

(ix) For each $a \neq 0$, there exists a scalar $a^{-1} \in F$, such that $aa^{-1} = a^{-1}a = 1$

(x) $(a + b) \cdot c = (ac) + (bc)$

It might seem like a daunting list but remember the following: a field is just a number system for which we can add, subtract, multiply, and divide following the usual commutative, associative, and distributive laws. The set of rational numbers, denoted $\mathbb{Q}$,[†] is a field. The set of real numbers, denoted $\mathbb{R}$, and the set of complex numbers, denoted $\mathbb{C}$, are also both fields. These are the fields that we are most familiar with. Note that the set of integers, denoted $\mathbb{Z}$,[††] is NOT a field, as the quotient of two integers is not always an integer, e.g. $\frac{1}{2} \notin \mathbb{Z}$. Likewise, the set of natural numbers, denoted $\mathbb{N}$, is NOT a field for the same reasons with division but also because the difference of two natural numbers is not always a natural number, e.g. $3 - 4 = -1 \notin \mathbb{N}$.

Fields are the number systems we are used to in previous algebra classes. For example, fields are exactly the environment where linear equations can be solved. Let F be a field and $ax + b = c$ is a linear equation with variable x and $a, b, c \in F$. Here the variable itself is a placeholder for some other scalar. A **solution** to this linear equation is an assignment to the variable x which makes the equation true. For linear equations, we see there is only one solution in any field:

$$\begin{aligned}
ax + b &= c && \\
(ax + b) + (-b) &= c + (-b) && \text{Existence of Additive Inverses } \textit{(iv)} \\
ax + (b - b) &= c - b && \text{Additive Associativity } \textit{(ii)} \\
ax + 0 &= c - b && \\
ax &= c - b && \text{Existence of Additive Identity } \textit{(iii)} \\
a^{-1}(ax) &= a^{-1}(c - b) && \text{Existence of Multiplicative Inverses } \textit{(ix)} \\
(a^{-1}a)x &= \frac{c-b}{a} && \text{Multiplicative Associativity } \textit{(vii)} \\
1x &= \frac{c-b}{a} && \\
x &= \frac{c-b}{a} && \text{Existence of Multiplicative Identity } \textit{(viii)}
\end{aligned}$$

Note that commutativity is used to guarantee that subtraction and division is well-defined. Therefore, a linear equation has a unique solution over a field.

Example 1.2.2. Solve the following equations:

(a) $2x + 1 = 6$ over $\mathbb{Q}$

$$2x + 1 = 6 \quad \Rightarrow \quad 2x = 6 - 1 = 5 \quad \Rightarrow \quad x = \boxed{\frac{5}{2}}.$$

(b) $(i + 1)x + (3 - i) = 4 + 5i$ over $\mathbb{C}$

$$\begin{aligned}
&(i + 1)x + (3 - i) = 4 + 5i \quad \Rightarrow \quad (i + 1)x = (4 + 5i) - (3 - i) = 1 + 6i \\
&\Rightarrow \quad x = \frac{1 + 6i}{1 + i} = \frac{1 + 6i}{1 + i}\left(\frac{1 - i}{1 - i}\right) = \frac{1 - i + 6i + 6}{1 - i + i + 1} = \boxed{\frac{7}{2} + \frac{5}{2}i}.
\end{aligned}$$

■

1.2.2 Modular Arithmetic and Finite Fields

We introduce now one last type of field: a finite field. Let n be a positive integer and let a be any integer. Consider the operation $a \div n$. The division algorithm known unto us since primary school guarantees there are unique integers q, r, called the *quotient* and *remainder*, respectively, such that $a = qn + r$ where $0 \leq r < n$. For example, for $n = 5$ and $a = 13$, we have that $13 = 2(5) + 3$, that is, $q = 2$ and $r = 3$. We generally would interpret this statement as 5 divides into 13 two times with remainder of three, that is, if 13 vectors are to be shared among 5 friends then each friend would get 2 vectors with 3 vectors left over.

Typically, with a division problem one is interested in finding the quotient but in **modular arithmetic**, one is instead interested in finding the remainder, which has many powerful uses. Let $\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$. Then $\mathbb{Z}_n$ should be viewed as the set of all possible remainders, or **residues**, when divided by n. Define next a function $\pmod n : \mathbb{Z} \to \mathbb{Z}_n$ by the rule $a \pmod n$ is the unique remainder of a when divided by n. For example, $8 \pmod 5 = 3$. Likewise, $14 \pmod 5 = 4$ since $14 = 2(5) + 4$, that is, 5 divides into 14 twice with a remainder of 4. The residue $a \pmod n$ is read "a modulo n" and n is called **modulus** of the function.

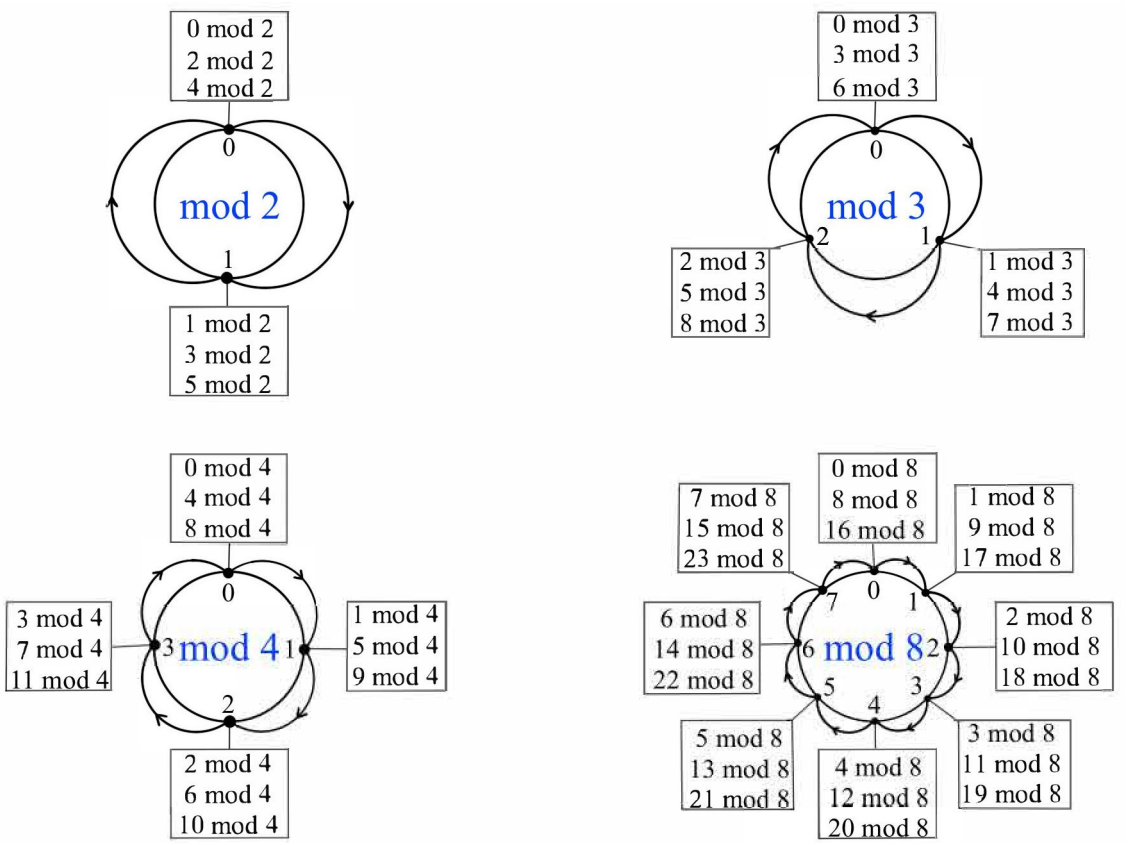
Note that if $0 \leq a < n$, then $a \pmod n = a$, the original value, for example, $3 \pmod 5 = 3$. In this example, the associated quotient is 0. This is a department from decimal division where the fraction $\frac{3}{5}$ may be expressed as the decimal 0.6. Special attention should also be given to when a is a negative integer. Note that $-11 \pmod 6 = 1$ since $-11 = (-2)(6) + 1$, where the quotient is -2 and the residue is 1. This is in contrast to the possibility $-11 = (-1)(6) + (-5)$. The difference here is that while quotients may be any integer: positive, zero, or negative, remainder must always exist in the range $0 \leq r \leq n-1$, which is necessarily nonnegative.

When $a \geq n$, one way to determine $a \pmod n$ is to subtract from a the modulus n until the difference lies in $\mathbb{Z}_n$. For example, $14 > 5$, the first difference $14 - (1)5 = 14 - 5 = 9$ is still too large. The second difference $14 - 2(5) = 9 - 5 = 4 \in \mathbb{Z}_5$. Thus, $14 \pmod 5 = 4$, as observed above. This notion of repeated subtraction is actually where the division algorithm gets its life, in a similar way of viewing multiplication as repeated addition. When a is negative, we can instead add the modulus n to a (or subtract a negative, if one prefers) until the sum lies inside of $\mathbb{Z}_n$. For examples, $-11 - (-1)6 = -11 + 6 = -5$ which is too small, but the next iteration $-11 - (-2)6 = -5 + 6 = 1$. Thus, $-11 \pmod 6 = 1$, as observed above.

Note that $7 \pmod 6 = 1 = -11 \pmod 6$, that is, it is possible for two different integers to have the same remainder when divided by the modulus n. We say that two integers a and b are **congruent modulo** n if $a \pmod n = b \pmod n$ and denote this congruence as $a \equiv b \pmod n$. For example, $7 \equiv 12 \pmod 5$ since 7 and 12 both have remainder 2. If $a \equiv b \pmod n$, then $a = q_1 n + r$ and $b = q_2 n + r$ for some quotients q_1, q_2. Then their difference $a - b = (q_1 - q_2)n$ is a multiple of the modulus n. In other words, $a \equiv b \pmod n$ if and only if $a - b$ is divisible by n. For example, $12 - 7 = 5$ which implies that $12 \equiv 7 \pmod 5$.

We define two operations on $\mathbb{Z}_n$ called **modular addition** and **modular multiplication**: modular addition is defined by the rule that two integers are added together and the sum's remainder modulo n is reported, modular multiplication is defined similarly. The final result should be a number in $\mathbb{Z}_n$.

A visual interpretation of modular addition would be the following. Imagine n pearls on a string labeled from left to right with the numbers 0 through $n-1$. Take the two ends of this string and connect them together to form a pearl necklace of integers. To compute $a + b \pmod n$, start at the pearl labeled a and rotate the necklace by one pearl to the right (counter-clockwise) exactly b times. The pearl you land on is then modular sum of a and b. For example, starting at 2 and counting forward on necklace with 5 pearls 4 times gives the pearl 1, that is, $2 + 4 \equiv 1 \pmod 5$. Modular multiplication can similarly be visualized as iterated addition on the pearl necklace.



Example 1.2.3. Let consider the following calculations:

(a) $6 + 11 \pmod 5$

$$6 + 11 \equiv 17 \equiv \boxed{2} \pmod 5$$

(b) $7 + 13 \pmod 5$

$$7 + 13 \equiv 20 \equiv \boxed{0} \pmod 5$$

(c) $2(4) \pmod 7$

$$2(4) \equiv 8 \equiv 1 \pmod 7.$$

(d) $2(5 + 6) \pmod{13}$

$$2(5 + 6) \equiv 2(11) \equiv 22 \equiv \boxed{9} \pmod{13}$$

(e) $3(5) + 8 \pmod{11}$

$$3(5) + 8 \equiv 15 + 8 \equiv 4 + 8 \equiv 12 \equiv \boxed{1} \pmod{11}$$ ∎

We can define **modular subtraction** $a - b \pmod n$ by adding the additive inverse of b with respect to modular addition, that is, $a - b \equiv a + (-b) \pmod n$. Thus, modular subtraction is the inverse operation of modular addition. Returning to the pearl necklace analogy above, if $a + b \pmod n$ turns the necklace to the right (counter-clockwise) b times from a, $a - b \pmod n$ turns the necklace to the left (clockwise) b times from a. For example, starting at 2 and counting backward on the necklace with 5 pearls 4 times gives the pearl 3, that is, $2 - 4 \equiv 3 \pmod 5$. Likewise, $14 - 5 \equiv 9 \equiv 1 \pmod 8$. Note that $-b \equiv n - b \pmod n$. Thus, the final result when computing modular subtraction should always be positive.

Similarly, we define **modular division** $a/b \pmod{n}$ by multiplying by the multiplicative inverse of b with respect to modular multiplication, that is, $a/b \equiv ab^{-1} \pmod{n}$. This multiplicative inverse will be an element of $\mathbb{Z}_n$ such that $bb^{-1} \equiv 1 \pmod{n}$, that is, their product will be one more than a multiple of n as this implies the remainder is 1 when divided by n (see Example 1.2.3.3). For example,

$$4^{-1} = \frac{1}{4} \equiv 2 \pmod{7}$$

since $4(2) = 8 \equiv 1 \pmod{7}$. Thus,

$$\frac{5}{4} \equiv 5(2) = 10 \equiv 3 \pmod{7}.$$

This multiplicative inverse does not always exists (see Exercise 1.2.2).

Alternatively, one can simplify a modular fraction without a multiplicative inverse by replacing the numerator with an integer equivalent to the original but actually divisible by the denominator in the usual sense. For example,

$$\frac{17}{3} \equiv \frac{17+10}{3} \equiv \frac{27}{3} \equiv \frac{9(3)}{3} \equiv 9 \pmod{5}.$$

The following theorem characterizes when this can be done.

Theorem 1.2.4. *The set $\mathbb{Z}_n$ is a field with respect to modular addition and modular multiplication if and only if n is a prime number.*

Example 1.2.5. Solve the linear equation $2x + 1 \equiv 6 \pmod{13}$.

Since $\mathbb{Z}_{13}$ is a field, we can solve this linear equation in the same fashion as the previous ones. Note

$$2x + 1 \equiv 6 \quad \Rightarrow \quad 2x \equiv 6 - 1 \equiv 5 \pmod{13} \quad \Rightarrow \quad x \equiv \frac{5}{2} \equiv 5(2)^{-1}.$$

Next, we need to either identify the reciprocal of $2 \pmod{13}$[‡] or replace 5 with a congruent integer which is even, that is, $\frac{5}{2} \equiv \frac{5+13}{2} \equiv \frac{18}{2} \equiv 9 \pmod{13}$. Therefore, the solution is 9. We can check this solution:

$$2(9) + 1 \equiv 18 + 1 \equiv 19 \equiv 6 \pmod{13}.$$ ∎

Example 1.2.6. Solve the linear equation $3x + 5 \equiv 1 \pmod{7}$.

We solve this equation similar to the previous example over the field $\mathbb{Z}_7$.

$$3x + 5 \equiv 1 \quad \Rightarrow \quad 3x \equiv 1 - 5 \equiv -4 \equiv 3 \pmod{7} \quad \Rightarrow \quad x \equiv \frac{3}{3} \equiv 1.$$

We can check this solution:

$$3(1) + 5 \equiv 3 + 5 \equiv 8 \equiv 1 \pmod{7}.$$ ∎

Exercises

(Go to Solutions)

1. Explain the difference between equality of integers and congruence modulo p.

For Exercises 2-6, determine with the statement is true or false. If false, correct the statement so that it is true.

2. The remainder $a \pmod{n}$ is read "a modulo n" and a is called the modulus of $\mathbb{Z}_n$.
3. It is possible for two different integers to have the same residue modulo n.
4. Modular division $a/b \pmod{n}$ is defined by multiplying by the inverse of b with respect to modular addition.
5. The set $\mathbb{Z}_n$ is a field with respect to modular addition and multiplication if and only if n is a real number.
6. If $0 \leq a < n$, then $a \pmod{n} = n$.

For Exercises 7-24, simplify the modular expression.

7. $-4 \pmod{7}$
8. $3(-7) \pmod{5}$
9. $-924 - 152 \pmod{4}$
10. $136 - 822 \pmod{13}$
11. $2237 \pmod{11}$
12. $\frac{1}{2} \mod 3$
13. $\frac{1}{3} \mod 7$
♠ 14. $34 + 61 \pmod{7}$
♠ 15. $4(13) \pmod{5}$
16. $5(15) \pmod{6}$
17. $2(8-5)+4 \pmod{3}$
♠ 18. $2(6-3)+7 \pmod{2}$
♠ 19. $3[2+1-(6+3)]+6(7) \pmod{13}$
♠ 20. $\frac{5(4)+3}{2(6)} \pmod{7}$
21. $\frac{3(5)+4}{2(7)} \pmod{9}$
22. $\frac{6(7)+5}{2(4)} \pmod{13}$
23. $2[(3+4+2+5)+1+2(3)-3+2(-2)] \pmod{7}$
24. $\frac{4(9)+1-(6+12)}{1+3(6)} \pmod{13}$
25. $\frac{-2(35-74)}{3} \pmod{19}$

For Exercises 26–43, solve the linear equation.

26. $7x + 10 = 3$
♠ 27. $6x + 7 = 2$
♠ 28. $(2+i)x + (5-7i) = 3 - 4i$
29. $3x + 7 \equiv 5 \pmod{5}$
♠ 30. $x + 1 \equiv 0 \pmod{2}$
♠ 31. $2x + 3 \equiv 4 \pmod{5}$
♠ 32. $3x + 2 \equiv 1 \pmod{7}$
♠ 33. $5x + 1 \equiv 4 \pmod{11}$
34. $420x - 7 \equiv 98 \pmod{11}$
35. $2x + 5 \equiv 0 \pmod{7}$
36. $14x + 12 \equiv 124 \pmod{6}$
37. $5x - 25 \equiv 71 \pmod{14}$
38. $2x - 4 \equiv 14 \pmod{7}$
39. $3(2-4(x+1)) \equiv 27 \pmod{3}$
40. $-20 - 4x \equiv 40 \pmod{4}$
41. $5x + 22 \equiv 15 \pmod{8}$
42. $6x + \frac{5}{2} \equiv 0 \pmod{7}$
43. $(3+10)x + 2(7) + 3 \equiv 5(14) \pmod{7}$

♠ 44. Show that the linear equation $2x + 3 \equiv 4 \pmod{6}$ has no solution.

---

†The symbol for the rational field $\mathbb{Q}$ comes from the English word ***Q****uotient*, given that every rational number can be expressed as a quotient of integers.

††The symbols for the ring of integers $\mathbb{Z}$ comes from the German word ***Z****ahlen*, which translates as *number*.

‡By the way, the reciprocal of 2 modulo 13 is 7 since $2(7) \equiv 14 \equiv 1 \pmod{13}$. Note that $x \equiv 5(2)^{-1} \equiv 5(7) \equiv 35 \equiv 9 \pmod{13}$.

"Your sacred space is where you can find yourself again and again." – Joseph Campbell

1.3 Vector Spaces

Now that we have learned about fields and linear equations, the next important linear structure should discuss is that of *linear combination*. In order to define this, we first need a notion of a vector. We say that a mathematical quantity is a *vector* if we can add two vectors together to create a third vector and we can scale a vector by a scalar (hence the name). Vectors are distinct from scalars. While a scalar is a number, a vector is more complicated than just a number. To distinguish the extra complexity of vectors, vector quantities are denoted either in boldface font, such as $\boldsymbol{u}, \boldsymbol{v}$, or with an arrowhead above the symbol, such as $\vec{u}, \vec{v}$. The following definition makes precise this notion.

Definition 1.3.1. A **vector space** over a field F is a nonempty set V of elements, called **vectors**, on which are defined two operations, called **addition** $+ : V \times V \to V$ and **scalar multiplication** $\cdot : F \times V \to V$, such that for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ and all scalars $c, d \in F$, the following eight axioms hold:

(i) **Additive Associativity**: $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$

(ii) **Additive Commutativity**: $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

(iii) **Additive Identity**: There exists a vector $\boldsymbol{0} \in V$, such that $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{0} + \boldsymbol{u} = \boldsymbol{u}$

(iv) **Additive Inverse**: For each $\boldsymbol{u}$, there exists a vector $-\boldsymbol{u} \in V$, such that $\boldsymbol{u} + (-\boldsymbol{u}) = (-\boldsymbol{u}) + \boldsymbol{u} = \boldsymbol{0}$

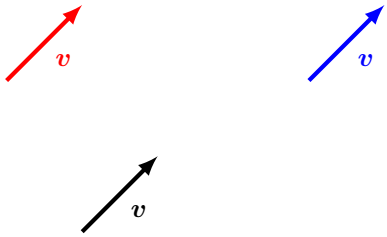
(v) **Left Distributivity**: $c(\boldsymbol{u} + \boldsymbol{v}) = c\boldsymbol{u} + c\boldsymbol{v}$

(vi) **Right Distributivity**: $(c + d)\boldsymbol{u} = c\boldsymbol{u} + d\boldsymbol{u}$

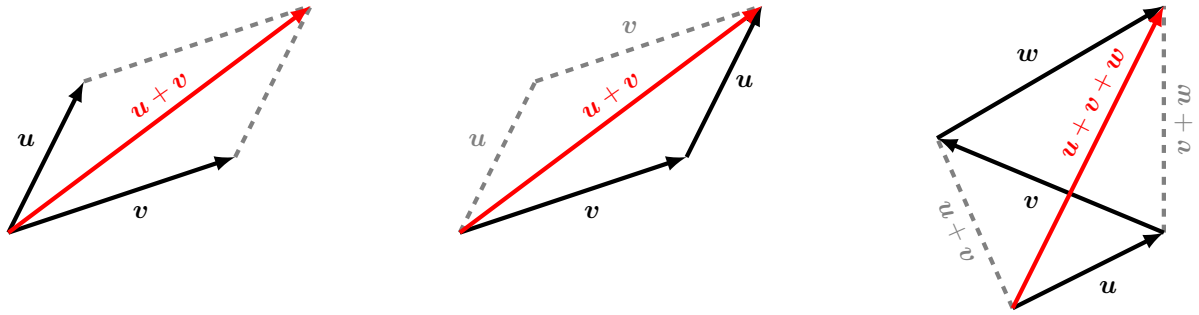
(vii) **Multiplicative Associativity**: $c(d\boldsymbol{u}) = (cd)\boldsymbol{u}$

(viii) **Multiplicative Identity**: $1\boldsymbol{u} = \boldsymbol{u}$

Example 1.3.2. In physics, a vector represents a mathematical quantity with both magnitude and direction, such as a force applied to an object. Vectors are thus represented as arrows pointing in the given direction and whose length represents the magnitude of the vector. Physical vectors are also free to move about in space, that is, this exact location does not matter so long as their length nor direction does not change. Therefore, the



three arrows illustrated above represent that same vector $\boldsymbol{v}$. Hence, the arrow is determined by the relative displacement of the head of the arrow from its tails. These vectors can be added together by the so-called *parallelogram rule*: begin by placing the tails of the two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ together, form a parallelogram whose parallel sides correspond to $\boldsymbol{u}$ with a copy of $\boldsymbol{u}$ and $\boldsymbol{v}$ with a copy of $\boldsymbol{v}$; the sum $\boldsymbol{u} + \boldsymbol{v}$ is the unique vector which is the diagonal of the parallelogram whose tail agrees with the tail of $\boldsymbol{u}$ and $\boldsymbol{v}$, as displayed:



The Parallelogram Rule $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$ $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$

The second diagram shows why arrow addition is commutative, as the two paths from the tail of $\boldsymbol{u} + \boldsymbol{v}$ to head produce the same vector. In particular, the black, solid path comprises $\boldsymbol{v} + \boldsymbol{u}$ while the gray, dashed

path comprises $\boldsymbol{u}+\boldsymbol{v}$, and both paths have the same starting and ending points. The third diagram likewise demonstrates how three or more vectors are added together and why arrow addition is associative, as the path following $(\boldsymbol{u}+\boldsymbol{v})$ then $\boldsymbol{w}$ produces the same vector as the path following $\boldsymbol{u}$ then $(\boldsymbol{v}+\boldsymbol{w})$.

The zero vector $\mathbf{0}$ would be simply a point in space, that is, the vector with no magnitude nor direction. Notice that adjoining a point to the head or tail will neither lengthen the arrow nor turn it. Thus, $\boldsymbol{v}+\mathbf{0}=\boldsymbol{v}$. Additionally, given a vector $\boldsymbol{v}$, $-\boldsymbol{v}$ is defined as the arrow with the same length but pointing in the opposite direction. It holds that $\boldsymbol{v}+(-\boldsymbol{v})=(-\boldsymbol{v})+\boldsymbol{v}=\mathbf{0}$.

We define scalar multiplication of arrows by multiplying the length of the arrow by the scalar (and switching directions if the scalar is negative). Clearly, $1\boldsymbol{v}=\boldsymbol{v}$, since the length of the arrow is unchanged, and $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$, since stretching a vector by a factor of d then stretching by a factor of c as the net affect of stretching the arrow by a factor of cd. It is left as an exercise of the reader to prove that two distributive laws. Therefore, the set of arrow in space forms a vector space under these operations. ∎

Example 1.3.3. Another type of vector is an array of numbers. Let F be a field. Then a **column vector** is an array of n scalars from F such that the array is oriented vertically, e.g. $\boldsymbol{v}=\begin{bmatrix}1\\2\\3\end{bmatrix}$. A **row vector** likewise is an array of n scalars such that the array is oriented horizontally, e.g. $\boldsymbol{v}=(1,2,3)$. Our notation of a row vector agrees with the commonly used notation for coordinates of points in Cartesian geometry. As such, we naturally identify these arrays of scalars with points in space. The difference between row vectors and column vectors is purely notational, and we will use the two notations interchangeably.† The order of the scalars in the array does matter, e.g. $(1,2,3)\neq(2,3,1)$.

Let F^n denote the set of all column vectors with n entries coming from the field F, for example, for example,

$$\begin{bmatrix}\pi\\0\\-\sqrt{17}\end{bmatrix}\in\mathbb{R}^3,\quad \begin{bmatrix}1-2i\\\frac{1}{2}-\frac{5}{2}i\end{bmatrix}\in\mathbb{C}^2,\quad \text{or}\quad \begin{bmatrix}1\\2\\0\\2\\1\end{bmatrix}\in\mathbb{Z}_3^5.\text{††}$$

Addition of column vectors is component-wise, that is, scalars in the corresponding positions are added together:

$$\boldsymbol{u}+\boldsymbol{v}=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}+\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}=\begin{bmatrix}u_1+v_1\\u_2+v_2\\\vdots\\u_n+v_n\end{bmatrix}. \tag{1.3.1}$$

For scalar multiplication, every component is multiplied by the scalar, that is:

$$c\boldsymbol{x} = c\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}. \tag{1.3.2}$$

Here the zero vector $\mathbf{0}$ is the array of all zeros and $-\boldsymbol{v} = (-1)\boldsymbol{v}$. We show that this definition of vector addition satisfies the additive associativity axiom:

$$\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \left(\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \right) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix} = \begin{bmatrix} u_1 + (v_1 + w_1) \\ u_2 + (v_2 + w_2) \\ \vdots \\ u_n + (v_n + w_n) \end{bmatrix}$$

$$\overset{(*)}{=} \begin{bmatrix} (u_1 + v_1) + w_1 \\ (u_2 + v_2) + w_2 \\ \vdots \\ (u_n + v_n) + w_n) \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \left(\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \right) + \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w},$$

where the marked equality $(*)$ follows from the additive associtivity of the field F. The remaining axioms of a vector space are left as an exercise to the reader. Therefore, F^n is a vector space, the one we will place the most focus on in this text. ∎

Example 1.3.4. Let $\boldsymbol{u} = \begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix} \in \mathbb{R}^3$. Then

$$\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} 6 \\ -2 \\ 2 \end{bmatrix} + \begin{bmatrix} -3 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} 6-3 \\ -2+0 \\ 2+5 \end{bmatrix} = \begin{bmatrix} 3 \\ -2 \\ 7 \end{bmatrix}.$$

If $\boldsymbol{v} = \begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} \in \mathbb{R}^3$, then

$$3\boldsymbol{v} = 3\begin{bmatrix} -1 \\ 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -3 \\ 9 \\ 6 \end{bmatrix}.$$ ∎

Definition 1.3.5. Given vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in F^m$ and scalars $c_1, c_2, \ldots, c_n \in F$, the vector $\boldsymbol{x}$ given as

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$$

is called a **linear combination** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ with **coefficients** $c_1, c_2, \ldots, c_n$.

A linear combination is a way of combining vectors using addition and scalar multiplication. These vectors operations, of course, depend entirely on the field of scalars.

Example 1.3.6. Let us simplify the following linear combinations over the vector spaces $\mathbb{C}^2$ and $\mathbb{Z}_3^3$, respectively.

(a) $$2\begin{bmatrix} 3 \\ i \end{bmatrix} + i\begin{bmatrix} 2+i \\ 3+5i \end{bmatrix} = \begin{bmatrix} 6 \\ 2i \end{bmatrix} + \begin{bmatrix} -1+2i \\ -5+3i \end{bmatrix} = \begin{bmatrix} 6+(-1+2i) \\ 2i+(-5+3i) \end{bmatrix} = \begin{bmatrix} 5+2i \\ -5+5i \end{bmatrix}.$$

(b) $$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \equiv \begin{bmatrix} 0+2+2 \\ 1+2+0 \\ 2+2+1 \end{bmatrix} \equiv \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \pmod 3.$$ ∎

Vectors can also be many other things, like matrices, functions, sequences of numbers. Even linear equations can be vectors.

Example 1.3.7. Let V be the set of linear equations with n variables, say $x_1, x_2, \ldots, x_n$ with coefficients coming from the field F. For example,

$$c_1x_1 + \ldots + c_nx_n = b_1$$

is an element of V for $c_1, \ldots, c_n, b_1 \in F$. Likewise, if

$$d_1x_1 + \ldots + d_nx_n = b_2$$

is another linear equation in V, then

$$\begin{array}{rrrrr} & (c_1x_1 & + \ldots + & c_nx_n & = & b_1) \\ + & (d_1x_1 & + \ldots + & d_nx_n & = & b_2) \\ \hline & (c_1+d_1)x_1 & + \ldots + & (c_n+d_n)x_n & = & (b_1+b_2) \end{array}$$

Note that $c_i + d_i, b_1 + b_2 \in F$ for all i. Thus, the sum of equations is a member of V, that is, it is a vector too. Similarly, if $a \in F$, then $c_1x_1 + \ldots + c_nx_n = b_1$ scaled by a is the linear equation

$$(ac_1)x_1 + \ldots + (ac_n)x_n = (ab_1).$$

As $ac_i, ab_1 \in F$ for all i, this equation is likewise a vector. Therefore, V is a vector space. Let $E_1, E_2, \ldots, E_m \in V$ be a list of linear equations with common solution $\boldsymbol{x} \in F^n$. Then $\boldsymbol{x}$ is a solution to any linear combination $a_1E_1 + a_2E_2 + \ldots + a_mE_m$. ∎

Theorem 1.3.8. *The following properties hold for any F-vector space V. Let $\boldsymbol{u} \in V$ and $c \in F$.*

(i) The zero vector $\mathbf{0}$ in V is unique.

(ii) The additive inverse of $\boldsymbol{u}$ is unique.

(iii) $0\boldsymbol{u} = \mathbf{0}$. *(iv)* $c\mathbf{0} = \mathbf{0}$. *(v)* $-\boldsymbol{u} = (-1)\boldsymbol{u}$.

Proof. *(i)* Suppose there is a vector $\boldsymbol{\theta} \in V$ such that $\boldsymbol{\theta} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{\theta} = \boldsymbol{v}$ for all $\boldsymbol{v} \in V$. Then

$$\mathbf{0} = \mathbf{0} + \boldsymbol{\theta} = \boldsymbol{\theta}.$$

Therefore, $\mathbf{0}$ is the unique vector with this property.

(ii) Like the last property, suppose that $\boldsymbol{u}' \in V$ has the property that $\boldsymbol{u} + \boldsymbol{u}' = \boldsymbol{u}' + \boldsymbol{u} = \mathbf{0}$. Then

$$\boldsymbol{u}' = \boldsymbol{u}' + \mathbf{0} = \boldsymbol{u}' + (\boldsymbol{u} + -\boldsymbol{u}) = (\boldsymbol{u}' + \boldsymbol{u}) + -\boldsymbol{u} = \mathbf{0} + -\boldsymbol{u} = -\boldsymbol{u}.$$

Therefore, $-\boldsymbol{u}$ is the unique vector with this property.

(iii) First,

$$0\boldsymbol{u} = (0 + 0)\boldsymbol{u} = 0\boldsymbol{u} + 0\boldsymbol{u}.$$

Adding $-0\boldsymbol{u}$ to both sides of the equation gives $\mathbf{0} = 0\boldsymbol{u}$.

The remaining parts are left as exercises to the reader. □

Exercises

(Go to Solutions)

For Exercises 1-1, determine with the statement is true or false. If false, correct the statement so that it is true.

1. For any vector space V, $\emptyset \in V$.

For Exercises 2-23, simplify the linear combination.

2. $5\begin{bmatrix}1\\2\\4\end{bmatrix}+3\begin{bmatrix}0\\2\\1\end{bmatrix}+4\begin{bmatrix}1\\1\\1\end{bmatrix}$

♠ 3. $3\begin{bmatrix}-1\\0\\1\end{bmatrix}+2\begin{bmatrix}0\\0\\1\end{bmatrix}-2\begin{bmatrix}3\\4\\0\end{bmatrix}$

4. $\begin{bmatrix}3\\2\\1\end{bmatrix}+\begin{bmatrix}5\\8\\10\end{bmatrix}+\begin{bmatrix}0\\0\\5\end{bmatrix}$

5. $-1\begin{bmatrix}3\\0\\1\end{bmatrix}+\begin{bmatrix}1\\2\\4\end{bmatrix}+3\begin{bmatrix}2\\4\\1\end{bmatrix}$

6. $7\begin{bmatrix}3\\0\\2\end{bmatrix}+2\begin{bmatrix}8\\-3\\5\end{bmatrix}-3\begin{bmatrix}-2\\-2\\0\end{bmatrix}$

7. $8\begin{bmatrix}1\\4\\3\end{bmatrix}+2\begin{bmatrix}5\\0\\4\end{bmatrix}-4\begin{bmatrix}1\\0\\1\end{bmatrix}$

8. $(2+5i)\begin{bmatrix}3\\4\end{bmatrix}-(3-2i)\begin{bmatrix}7+i\\6+2i\end{bmatrix}$

♠ 9. $(3+2i)\begin{bmatrix}1\\2\end{bmatrix}-(1-2i)\begin{bmatrix}3+4i\\1+i\end{bmatrix}$

10. $(2+i)\begin{bmatrix}1+i\\i\end{bmatrix}+\begin{bmatrix}2i\\1+2i\end{bmatrix}$

11. $(2+4i)\begin{bmatrix}2\\3\end{bmatrix}-(3-2i)\begin{bmatrix}1+2i\\2-i\end{bmatrix}$

12. $(1+i)\begin{bmatrix}1-i\\2+i\end{bmatrix}-(1-i)\begin{bmatrix}3+i\\4-i\end{bmatrix}$

13. $(1-i)\begin{bmatrix}1\\2\\-1\end{bmatrix}+\begin{bmatrix}2\\3\\5\end{bmatrix}-(2+3i)\begin{bmatrix}1\\4\\2\end{bmatrix}$

14. $(4+5i)\begin{bmatrix}3\\5\\2\end{bmatrix}+(3-2i)\begin{bmatrix}1+3i\\2-3i\\2+5i\end{bmatrix}$

15. $2\begin{bmatrix}2\\1\end{bmatrix}-\begin{bmatrix}1\\0\end{bmatrix}+2\begin{bmatrix}1\\1\end{bmatrix}\pmod 3$

16. $-10\begin{bmatrix}3\\2\end{bmatrix}+\begin{bmatrix}5\\6\end{bmatrix}\pmod 7$

♠ 17. $\begin{bmatrix}1\\0\\1\end{bmatrix}+\begin{bmatrix}1\\1\\0\end{bmatrix}+\begin{bmatrix}1\\1\\1\end{bmatrix}\pmod 2$

18. $6\begin{bmatrix}1\\3\\4\end{bmatrix}+2\begin{bmatrix}1\\1\\5\end{bmatrix}+\begin{bmatrix}1\\1\\1\end{bmatrix}\pmod 3$

19. $\begin{bmatrix}3\\4\\3\end{bmatrix}+\begin{bmatrix}4\\1\\1\end{bmatrix}+\begin{bmatrix}0\\1\\1\end{bmatrix}\pmod 5$

20. $2\begin{bmatrix}1\\2\\3\end{bmatrix}+3\begin{bmatrix}3\\2\\1\end{bmatrix}+4\begin{bmatrix}0\\1\\0\end{bmatrix} \pmod 5$

♠ 21. $4\begin{bmatrix}1\\0\\0\\2\end{bmatrix}+2\begin{bmatrix}2\\3\\1\\0\end{bmatrix}+\begin{bmatrix}1\\2\\3\\4\end{bmatrix} \pmod 5$

22. $3\begin{bmatrix}-1\\2\\-3\\0\end{bmatrix}-2\begin{bmatrix}1\\2\\2\\1\end{bmatrix}+\begin{bmatrix}2\\3\\0\\1\end{bmatrix} \pmod 5$

23. $\begin{bmatrix}4\\2\\3\\1\end{bmatrix}+2\begin{bmatrix}0\\1\\4\\3\end{bmatrix}+3\begin{bmatrix}1\\1\\2\\3\end{bmatrix} \pmod 5$

♠ 24. Prove parts (iv) and (v) of Theorem 1.3.8.

♠ 25. Let $\mathcal{P}_n = \{a_0 + a_1x + a_2x^2 + \ldots + a_nx^n \mid a_0, a_1, a_2, \ldots, a_n \in F\}$ be the set of polynomials of degree at most n and coefficients from a field F. Show that $\mathcal{P}_n$ is a vector space, hence we may view polynomials as vector quantities. (More specifically, show that the sum of two polynomials is again a polynomial and the scalar multiple of a polynomial is likewise a polynomial, similar to the method introduced in Example 1.3.7).

Note that the way we defined vector addition and scalar multiplication over F^n is essentially the only way to do it to guarantee the vector space axioms. For Exercises 26-29, we will redefine scalar multiplication over F^3 (and keep vector addition the same). Find a counterexample for why this alternative definition of scalar multiplication fails at least one vector space axiom or property and hence does not produce a genuine vector space. Answers may vary.

♠ 26. Redefine scaling: $c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}$

♠ 27. Redefine scaling: $c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}cx_1\\x_2\\x_3\end{bmatrix}$

♠ 28. Redefine scaling: $c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}-cx_1\\-cx_2\\-cx_3\end{bmatrix}$

♠ 29. Redefine scaling: $c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}0\\cx_2\\cx_3\end{bmatrix}$

†From a vector perspective, the notion of a column vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, a row vector [1 2 3], and a geometric point $(1, 2, 3)$ are all the same. Why have different notations then? In other contexts, it is helpful to distinguish between them. In calculus or physics, a distinction between points and vectors (aka arrows) is sometimes desired (although often unnecessary). Thus, points have their usual notation and arrows are given a different notation, which may include the column or row vector notation we have already introduced or some other notation that we will not use here, e.g. $\langle 1, 2, 3 \rangle$. As arrows can be moved anywhere in space without changing their value, the vectors are often placed in *standard position*, meaning its tail is on the origin. Then the head of the arrow is a unique point in space which characterizes this arrow. This standard representation of the vector, sometimes called its *algebraic form*, is very desirable as the algebraic computations are far simpler their their geometric counterparts. For this reason, we see not strong reason to distinguish between arrows in space and the coordinates that they are pointing at. The only perspective that we will distinguish between a column vector and a row vector herein is when we consider them matrices, such as in Chapter (ii). This may be imperative for matrix operations, such as multiplication. Under the matrix perspective, a column vector is just an $n \times 1$ matrix, and a row vector is just a $1 \times n$ matrix.

††The combined notations of the set $\mathbb{Z}_p$, which is a field whose subscript defines the modulus, and the vector space F^n, whose superscripts defines the number of entries in each vector, conveys information about the vector space $\mathbb{Z}_p^n$. For example, $\mathbb{Z}_3^5$ denotes the vector space whose vectors comprise of five elements in the array that are integers congruent modulo 3, e.g, $\begin{bmatrix} 1 \\ 2 \\ 0 \\ 2 \\ 1 \end{bmatrix}$. Note that no integer in $\mathbb{Z}_3^n$ ever needs to be greater that 2 or smaller than 0, as all those integers are reducible modulo 3. Hence, $\mathbb{Z}_3^5$ contains exactly $3^5 = 243$ distinct vectors. In general, the vector space $\mathbb{Z}_p^n$ will contain exactly p^n many vectors.

"Don't spend time beating on a wall, hoping to transform it into a door." – Coco Chanel

1.4 Linear Transformations

The next important linear structure we have seen before (primarily in calculus) is the notion of a *linear operator* or a **linear transformation**. This is a function on vectors which preserve linear combinations.

Definition 1.4.1. Let X and Y be vector space over a field F. A function $T : X \to Y$ is a **linear transformation** if:

(i) $T(\boldsymbol{u} + \boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$, for all vectors $\boldsymbol{u}, \boldsymbol{v} \in X$;

(ii) $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$, for all vectors $\forall \boldsymbol{u} \in X$ and scalars $c \in F$.

Recall that X and Y are called the **domain** and **codomain** of T, respectively. For any $\boldsymbol{x} \in X$, we call $T(\boldsymbol{x})$ the **image** of $T\boldsymbol{x}$, and we call $\text{im}(T) = \{T(\boldsymbol{x}) \mid \boldsymbol{x} \in X\}$ (the set of images) the **image** (or **range**) of T. The **kernel** of a linear transformation T is the set of vectors which map to the zero vector, that is, $\ker T = \{\boldsymbol{x} \mid T(\boldsymbol{x}) = \boldsymbol{0}\}$, denoted $\ker(T)$.

Example 1.4.2. In calculus, there are many linear operators. For example, the derivative is a linear transformation since

$$\frac{d}{dx}[f(x) + g(x)] = \frac{d}{dx}[f(x)] + \frac{d}{dx}[g(x)] \qquad \text{and} \qquad \frac{d}{dx}[cf(x)] = c\frac{d}{dx}[f(x)].$$

Note that $\ker\left(\dfrac{d}{dx}\right)$ is the set of constant functions. Likewise, limits are linear operators:

$$\lim_{x\to a}[f(x) + g(x)] = \lim_{x\to a}[f(x)] + \lim_{x\to a}[g(x)] \qquad \text{and} \qquad \lim_{x\to a}[cf(x)] = c\lim_{x\to a}[f(x)].$$

Likewise, antiderivatives, definite integrals, indefinite integrals, and series are all linear operators on functions. Linear algebra was everywhere in calculus, we just didn't know it! ∎

Proposition 1.4.3. *Let $T : X \to Y$ be a linear transformation. Let $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in X$ and $c_1, \ldots, c_n \in F$. Then*

$$T(c_1\boldsymbol{x}_1 + \ldots + c_n\boldsymbol{x}_n) = c_1T(\boldsymbol{x}_1) + \ldots + c_nT(\boldsymbol{x}_n).$$

The converse also is true, that is, T is a linear transformation if it preserves linear combinations. In particular, $T(\boldsymbol{0}) = \boldsymbol{0}$.

Example 1.4.4. Consider the function $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by the rule $T(x_1, x_2, x_3) = (x_1 + 2x_2, x_3 - 3x_2)$, or in vector notation as:

$$T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ x_3 - 3x_2 \end{bmatrix}.$$

This is a function which takes a vector in 3-space, such as $(1, 2, 3)$, and maps it onto a vector in 2-space, namely $T(1, 2, 3) = (1 + 2(2), 3 - 3(2)) = (5, -3)$. But more than just a function that sends 3D vectors onto 2D vectors, this function *preserves* the structure of the vector space $\mathbb{R}^3$ as it maps onto $\mathbb{R}^2$. This function is a linear transformation since

$$T\left(a\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + b\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right) = T\begin{bmatrix} ax_1 + by_1 \\ ax_2 + by_2 \\ ax_3 + by_3 \end{bmatrix} = \begin{bmatrix} (ax_1 + by_1) + 2(ax_2 + by_2) \\ (ax_3 + by_3) - 3(ax_2 + by_2) \end{bmatrix}$$

$$a\begin{bmatrix} x_1+2x_2 \\ x_3-3x_2 \end{bmatrix} + b\begin{bmatrix} y_1+2y_2 \\ y_3-3y_2 \end{bmatrix} = aT\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + bT\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}.$$

We next compute the kernel of T. Suppose $T(\boldsymbol{x}) = \mathbf{0}$. This implies that $\begin{bmatrix} x_1+2x_2 \\ x_3-3x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Comparing components, we find two linear equations $x_1 + 2x_2 = 0$ and $x_3 - 3x_2 = 0$. Together, this forms a homogeneous system of linear equations.

$$\begin{cases} x_1 + 2x_2 = 0 \\ -3x_2 + x_3 = 0. \end{cases}$$

If we attempt to solve this system, we could use the method of substitution. Note we can solve for x_1 in the first equation, which gives $x_1 = -2x_2$. Likewise, we can solve x_3 in the second equation and find $x_3 = 3x_2$. What we see now is that x_1 and x_3 are determined by our choice of x_2 and there appears to be NO restriction on our choice of x_2, that is, we may choose it freely. For example, if $x_2 = 1$, it would mean that $x_1 = -2$ and $x_3 = 3$, that is, x_1 and x_3 are dependent on x_2. Note that

$$T\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} (-2)+2(1) \\ (3)-3(1) \end{bmatrix} = \begin{bmatrix} -2+2 \\ 3-3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \mathbf{0}.$$

Thus, this vector is in the kernel. In fact, the solution set to this system is exactly the kernel of T. If we let $x_2 = t$, some free parameter t, then we see that

$$\ker T = \left\{ \begin{bmatrix} -2t \\ t \\ 3t \end{bmatrix} \middle| \, t \in \mathbb{R} \right\} = \left\{ t\begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} \middle| \, t \in \mathbb{R} \right\} = \operatorname{Span}\left\{ \begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix} \right\}.$$

Is the vector $\boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ in the image of T? If so, there would be a vector $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$. But this implies that $\begin{bmatrix} x_1+2x_2 \\ x_3-3x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Again, this vector equation implies two linear equations, namely, $x_1 + 2x_2 = 1$ and $-3x_2 + x_3 = 2$, which as a system of equations is

$$\begin{cases} x_1 + 2x_2 = 1 \\ -3x_2 + x_3 = 2. \end{cases}$$

Again by substitution, we see $x_1 = 1 - 2x_2$ and $x_3 = 2 + 3x_2$. Thus, if $x_2 = 0$, we get that

$$T\left(\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1+2(0) \\ 2-3(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Therefore, yes, $\boldsymbol{b} \in \operatorname{im} T$. Of course, many other choices of $\boldsymbol{x}$ were possible such that $T(\boldsymbol{x}) = \boldsymbol{b}$. ∎

Notice that the components of $T(\boldsymbol{x})$ in the previous example are linear combinations of the components of the input vectors, that is, each slot in $T(\boldsymbol{x})$ is a linear combination of the variables $x_1, x_2, \ldots$ It can be shown that a function $T : F^n \to F^m$ is a linear transformation if and only if the components of $T(\boldsymbol{x})$ are linear combinations of the components of $\boldsymbol{x}$. Another takeaway from the previous example is that computing the kernel or image of a linear transformation resulted in solving a system of linear equations. This is no coincidence and we will discuss solving systems of linear equations more in the next sections.

Example 1.4.5. Consider the function $T : \mathcal{P}_2 \to \mathbb{R}^2$ given by the rule $T(ax^2 + bx + c) = \begin{bmatrix} 5a + 2c \\ -a + c \end{bmatrix}$.

Let $f(x) = a_1x^2 + b_1x + c_1, g(x) = a_2x^2 + b_2x + c_2 \in \mathcal{P}_2$ and $r \in \mathbb{R}$. Then

$$T(f(x) + g(x)) = T((a_1x^2 + b_1x + c_1) + (a_2x^2 + b_2x + c_2)) = T((a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2))$$
$$= \begin{bmatrix} 5(a_1 + a_2) + 2(c_1 + c_2) \\ -(a_1 + a_2) + (c_1 + c_2) \end{bmatrix} = \begin{bmatrix} (5a_1 + 2c_1) + (5a_2 + 2c_2) \\ (-a_1 + c_1) + (-a_2 + c_2) \end{bmatrix} = \begin{bmatrix} 5a_1 + 2c_1 \\ -a_1 + c_1 \end{bmatrix} + \begin{bmatrix} 5a_2 + 2c_2 \\ -a_2 + c_2 \end{bmatrix}$$
$$= T(a_1x^2 + b_1x + c_1) + T(a_2x^2 + b_2x + c_2) = T(f(x)) + T(g(x))$$

and

$$T(rf(x)) = T(r(a_1x^2 + b_1x + c_1)) = T((ra_1)x^2 + (rb_1)x + (rc_1))$$
$$= \begin{bmatrix} 5(ra_1) + 2(rc_1) \\ -(ra_1) + (rc_1) \end{bmatrix} = \begin{bmatrix} r(5a_1 + 2c_1) \\ r(-a_1 + c_1) \end{bmatrix} = r \begin{bmatrix} 5a_1 + 2c_1 \\ -a_1 + c_1 \end{bmatrix} =$$
$$rT(a_1x^2 + b_1x + c_1) = rT(f(x)).$$

Therefore we see that T is a linear transformation between the vector spaces $\mathcal{P}_2$ and $\mathbb{R}^2$.

To find the kernel of T, we must solve the vector equation $\begin{bmatrix} 5a + 2c \\ -a + c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, which corresponds to the linear system $\begin{cases} 5a + 2c = 0 \\ -a + c = 0. \end{cases}$ We solve this system by elimination. Scale the second equation by 5 and add together the resulting equation with the first equation. This gives $7c = 0$, which implies that $c = 0$. The original second equation $-a + c = 0$ implies that $a = c$. Hence, $a = 0$ likewise. What about the coefficient b? The map T seems to "forget" about the coefficient b, that is, the calculation of the image $T(ax^2 + bx + c)$ is independent of the coefficient b. The above linear system is actually a system with two equations and three unknowns, although none of the equations place any restriction on the coefficient b. Hence, $a = c = 0$ are dependent variables and b is a free variable. Therefore, $\ker T = \{bx \mid b \in \mathbb{R}\} \subseteq \mathcal{P}_2$.

To find the image of T, we consider the general vector equation $\begin{bmatrix} 5a + 2c \\ -a + c \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$, which corresponds to the linear system $\begin{cases} 5a + 2c = x \\ -a + c = y. \end{cases}$ Reducing this linear system by the same elimination technique as we did with the kernel, we get $7c = x + 5y$ or $c = \dfrac{x + 5y}{7}$. Since $a = c - y$, we also get that $a = \dfrac{x + 5y}{7} - y = \dfrac{x - 2y}{7}$. For example, if we wanted a polynomial which mapped onto $\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ via T, we could select $f(x) = 4x^2 + 3$ (like we saw in the kernel calculation, the coefficient b is meaningless via this transformation). In particular, $\operatorname{im} T = \mathbb{R}^2$. ∎

Definition 1.4.6. A mapping $T : X \to Y$ is said to be **onto** (or **surjective**) if for each $\boldsymbol{b} \in Y$ there is at least one $\boldsymbol{x} \in X$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$, that is, the codomain and range of T are the same.

A mapping $T : X \to Y$ is said to be **one-to-one** (or **injective**) if for each $\boldsymbol{b} \in Y$ there is at most one $\boldsymbol{x} \in X$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$, that is, $T(\boldsymbol{u}) = T(\boldsymbol{v})$ implies that $\boldsymbol{u} = \boldsymbol{v}$.

A mapping $T : X \to Y$ is said to be **invertible** (or **bijective**) if for each $\boldsymbol{b} \in Y$ there is exactly one $\boldsymbol{x} \in X$, that is, T is one-to-one and onto. In particular, there exists another linear transformation $S : Y \to X$ such that $S \circ T = \mathrm{Id}$ and $T \circ S = \mathrm{Id}$, where Id is the identity function $\boldsymbol{x} \mapsto \boldsymbol{x}$.

Proposition 1.4.7. *Let $T : X \to Y$ be a linear transformation. Then T is one-to-one if and only if* $\ker T = \{\mathbf{0}\}$. *Additionally, T is onto if and only if* $\operatorname{im} T = Y$.

Example 1.4.8. In Example 1.4.4, we see that T is not one-to-one since the kernel is nontrivial, that is, both $\mathbf{0}$ and $\begin{bmatrix} 2 \\ -1 \\ -3 \end{bmatrix}$ have the same image.

We can determine whether T is onto by determining if a generic vector in $\mathbb{R}^2$, say $\boldsymbol{b} = (b_1, b_2)$ is an image of some vector $\boldsymbol{x} \in \mathbb{R}^3$. Now this becomes a vector equation:

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} x_1 + 2x_2 \\ x_3 - 3x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \quad \Rightarrow \quad \begin{cases} x_1 + 2x_2 & = b_1 \\ -3x_2 + x_3 & = b_2 \end{cases}.$$

Treating x_2 as a free variable and setting it equal to 0, we see that $x_1 = b_1$ and $x_3 = b_2$. Therefore, $T(b_1, 0, b_2) = (b_1, b_2) = \boldsymbol{b}$. As $\boldsymbol{b}$ was no specific vector in $\mathbb{R}^2$, but, in fact, a generic vector, this shows that any vector in $\mathbb{R}^2$ can be the image of a vector from $\mathbb{R}^3$ via T, that is, T is onto. ∎

Example 1.4.9. Consider the linear transformation $T : \mathbb{Z}_2^3 \to \mathbb{Z}_2^4$ given by the rule:

$$(x_1, x_2, x_3) \mapsto (x_1, x_2, x_3, x_1 + x_2 + x_3).$$

This time, if $T(x_1, x_2, x_3) = (0, 0, 0, 0)$ then $x_1 = x_2 = x_3 = 0$. Thus, $\ker T = \{(0, 0, 0)\}$. Therefore, T is one-to-one. On the other hand, consider $\boldsymbol{b} = (1, 1, 1, 0)$. If $\boldsymbol{b} \in \operatorname{im} T$, then $x_1 = x_2 = x_3 = 1$. But $x_1 + x_2 + x_3 \equiv 1 \not\equiv 0$. Thus, $\boldsymbol{b} \neq T(\boldsymbol{x})$ for any $\boldsymbol{x} \in \mathbb{Z}_2^3$. Therefore, T is not onto. Note this linear transformation determines an error detecting code, because if an error occurred in any of the first three bits then the sum of the first three transmitted bits would not add up to the last bit. ∎

Exercises

(Go to Solutions)

For Exercises 1-5, use the transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by the rule:

$$T(x, y) = (x + y, 0, 2x + 3y).$$

♠ 1. Show that T is a linear transformation.

♠ 2. Compute $T(1, 2)$ and $T(5, -2)$.

♠ 3. Compute $\ker T$.

♠ 4. Is $\boldsymbol{b} = (1, 0, 1)$ in the image of T? Explain.

♠ 5. Determine whether T is one-to-one, onto, or neither.

For Exercises 6-10, use the transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by the rule:

$$T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_3 \\ x_1 + 3x_2 \end{bmatrix}.$$

6. Show that T is a linear transformation.

7. Compute $T\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ and $T\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$.

8. Compute $\ker T$.

9. Is $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in the image of T? Explain.

10. Determine whether T is one-to-one, onto, or neither.

For Exercises 11-15, use the transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by the rule:

$$T(x, y, z) = (x + y - 2z, -y + z).$$

11. Show that T is a linear transformation.

♠ 12. Compute $T(1, 2, 3)$ and $T(1, 0, -2)$.

♠ 13. Compute $\ker T$.

♠ 14. Is $\boldsymbol{b} = (3, -1)$ in the image of T? Explain.

♠ 15. Determine whether T is one-to-one, onto, or neither.

For Exercises 16-18, use the transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ given by the rule:

$$T(x, y, z) = (2z + y, 2x + 4, -2y).$$

16. Compute $T(1, 3, 2)$.

17. Is $\boldsymbol{b} = (1, 6, 2)$ in the image of T? Explain.

18. Show that T is NOT a linear transformation.

For Exercises 19-21, use the transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ given by the rule:

$$T(x, y, z) = (2, 5).$$

19. Show that T is NOT a linear transformation.

20. Is $\boldsymbol{b} = (3, 2)$ in the image of T? Explain.

21. Determine whether T is one-to-one, onto, or neither.

For Exercises 22-26, use the transformation $T : \mathbb{Z}_2^4 \to \mathbb{Z}_2$ given by the rule:

$$T(x, y, z, w) \equiv x + y + z + w.$$

22. Show that T is a linear transformation.

♠ 23. Compute $T(1, 0, 0, 1)$ and $T(1, 0, 1, 1)$.

♠ 24. Compute $\ker T$.

♠ 25. Is $\boldsymbol{b} = 1$ in the image of T? Explain.

♠ 26. Determine whether T is one-to-one, onto, or neither.

For Exercises 27-27, use the transformation $T : \mathbb{Z}_3^2 \to \mathbb{Z}_3$ given by the rule:

$$T(x, y) = 2x + y.$$

27. Compute $T(1, 2)$ and $T(5, 2)$.

For Exercises 28-29, use the transformation $T : \mathbb{Z}_5^3 \to \mathbb{Z}_5^3$ given by the rule:

$$T(x, y, z) = \begin{bmatrix} x + 2y + z \\ 2x + 2y + 2z \\ 3z + 4y + 3z \end{bmatrix}.$$

28. Compute $T(1, 0, 1)$ and $T(1, 2, 3)$.

29. Is $\boldsymbol{b} = (2, 4, 3)$ in the image of T? Explain.

For Exercises 30-32, use the transformation $T : F^n \to F$ given by the rule:

$$T(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \ldots + a_nx_n,$$

where F is a field and $a_1, a_2, \ldots, a_n \in F$.

30. Show that T is a linear transformation.

31. Is it possible to determine if T is one-to-one with this information? Why or why not?

32. Is it possible to determine if T is onto with this information? Why or why not?

"When you put your hand to the plow, you can't put it down until you get to the end of the row."
– Alice Paul

1.5 Augmented Matrices

/Let us begin to find an effective algorithm for solving linear systems. The following three *Elementary Row Operations*† will be the basic techniques used to solve linear systems. In the elementary row operations, we refer to equations as "rows" for reasons that will be more clear by the end of this section.

Elementary Row Operations

1. (Replacement) Replace one row by the sum of itself and a multiple of another row.

2. (Interchange) Interchange the order of any two rows in the system.

3. (Scaling) Multiply all scalars in a row by a nonzero scalar.

We say that two systems of linear equations are **row equivalent** if there is a sequence of row operations that transforms one system into the other.

Theorem 1.5.1. *Two linear systems are equivalent if and only if they are row equivalent.*

Example 1.5.2. Solve

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 0 \\ 2x_2 - 8x_3 = -8 \\ -4x_1 + 6x_2 + 2x_3 = 10 \end{cases}$$

using row operations.

We begin by replacing Row 3 with (Row 3 + 4 ∗ Row 1). This gives:

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 0 \\ 2x_2 - 8x_3 = -8 \\ -2x_2 + 10x_3 = 10 \end{cases}$$

Next, we scale Row 2 by a factor of $\frac{1}{2}$, which gives:

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 0 \\ x_2 - 4x_3 = -4 \\ -2x_2 + 10x_3 = 10 \end{cases}$$

Lastly, we replace Row 3 with (Row 3 + 2 ∗ Row 2), which gives:

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 0 \\ x_2 - 4x_3 = -4 \\ 2x_3 = 2 \end{cases}$$

At this point, we recognize that we have solved for x_3, that is, $x_3 = 1$. Substituting this into Row 2 gives

$$x_2 - 4x_3 = -4 \quad \Rightarrow \quad x_2 = -4 + 4x_3 = -4 + 4(1) = -4 + 4 = 0.$$

Now we have solved for x_2. Lastly, we use these values to solve for x_1:

$$x_1 - 2x_2 + 2x_3 = 0 \quad \Rightarrow \quad x_1 = 2x_2 - 2x_3 = 2(0) - 2(1) = 0 - 2 = -2.$$

Therefore, $(x_1, x_2, x_3) = \boxed{(-2, 0, 1)}$ is the unique solution to the above system. ∎

Definition 1.5.3. If m and n are positive integers, an $m \times n$ **matrix** is a rectangular array of scalars with m rows and n columns. Note that the numbers of rows always comes first.

For example,

$$\begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & -3 \end{bmatrix}$$

is a 2×3 matrix over $\mathbb{R}$.

The essential information of a linear system can be recorded compactly in a matrix. Consider the system from the previous example

$$\begin{cases} x_1 - 2x_2 + 2x_3 = 0 \\ 2x_2 - 8x_3 = -8 \\ -4x_1 + 6x_2 + 2x_3 = 10 \end{cases}$$

Aligning similar variables into columns, we create the **coefficient matrix**

$$\begin{bmatrix} 1 & -2 & 2 \\ 0 & 2 & -8 \\ -4 & 6 & 2 \end{bmatrix}$$

and the **augmented matrix**

$$\left[\begin{array}{ccc|c} 1 & -2 & 2 & 0 \\ 0 & 2 & -8 & -8 \\ -4 & 6 & 2 & 10 \end{array}\right]$$

Example 1.5.4. Solve the following system of equations:

$$\begin{cases} 3x_2 + 3x_3 = 11 \\ 2x_1 - 3x_2 + 3x_3 = -4 \\ x_1 + x_2 + 4x_3 = 3 \end{cases}$$

using the augmented matrix.

The augmented matrix is:

$$\left[\begin{array}{ccc|c} 0 & 3 & 3 & 11 \\ 2 & -3 & 3 & -4 \\ 1 & 1 & 4 & 3 \end{array}\right]$$

Now, the row operations for linear systems work on augmented matrices as well. Interchanging Row's 1 and 3 gives:

$$\left[\begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 2 & -3 & 3 & -4 \\ 0 & 3 & 3 & 11 \end{array}\right]$$

Next, replace Row 2 with (Row 2 − 2Row 1), which gives:

$$\left[\begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 0 & -5 & -5 & -10 \\ 0 & 3 & 3 & 11 \end{array}\right]$$

Next, scale Row 2 by $-\frac{1}{5}$ to get:

$$\left[\begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 3 & 3 & 11 \end{array}\right]$$

Finally, replace Row 3 with (Row 3 − 3Row 2), which gives:

$$\left[\begin{array}{ccc|c} 1 & 1 & 4 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 5 \end{array}\right]$$

This implies that

$$\begin{cases} x_1 + x_2 + 4x_3 = 3 \\ \quad x_2 + x_3 = 2 \\ \quad\quad 0 = 5 \end{cases}$$

which is impossible. Therefore, the system has no solution, that is, the system is $\boxed{\text{inconsistent}}$. ∎

Definition 1.5.5. In a matrix, we say a row is a **zero row** if all scalars in this row are zero. Otherwise, we call it a **nonzero row**. In a nonzero row, we say the **leading entry** of the row is the leftmost, nonzero entry in the row.

A matrix is in (**row**) **echelon form** if it has the following three properties:

1. There is no zero row above a nonzero row.
2. Each leading entry is in a column to the right of the leading entry of the row above it.
3. All entries in a column below a leading entry are zero.

Essentially, 1. requires all zero rows to be at the bottom of a matrix in echelon form. Additionally, 2. and 3. require there exists a downward staircase of zeros in the lower left of the matrix, hence the name echelon.

A matrix in echelon form is in (**row**) **reduced echelon form** (or **RREF**) if additionally:

4. The leading entry in each nonzero row is 1.
5. All entries in a column above a leading entry are zero.

When considering whether an augmented matrix is in (row reduced) echelon form, consider only those columns to the left of the vertical line. That is, an augmented matrix is in (row reduced) echelon form if and only if its coefficient matrix is.

A **pivot position**, or simply just a **pivot**, in a matrix is a location that corresponds to a leading entry in one of its echelon forms. A **pivot column** (or **row**) is a column (or row) that contains a pivot position. The number of pivots of a matrix A is called its **rank**, denoted $\text{rank}(A)$.

A priori, we do not know the locations of the pivots in a matrix, but we can easily see them when the matrix is in echelon form. Now if two matrices are row equivalent, then their pivot positions will be the same. Thus, it will be highly useful to be able to compute echelon forms row equivalent to given matrices.

Example 1.5.6. The following two matrices are in echelon form:

$$\left[\begin{array}{ccc|c} \boxed{1} & i & 2-5i & 3 \\ 0 & 0 & \boxed{1} & 2-i \\ 0 & 0 & 0 & 5 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{ccc|c} \boxed{1} & 0 & 0 & -3 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 4 \end{array}\right].$$

The first matrix is NOT in row reduced echelon form, since there are nonzero entries above the pivot in the third column. The second one is and the first three columns are pivot columns. The first matrix has rank 2, and the second has rank 3. ∎

Solving systems of linear equations when in echelon form is very simple. Solve first for the equation involving one variable. Substitute this assignment of the variable into the linear equation involving two variables. Solve for the remaining unknown. Plug these two assignments into the equation with three unknowns. Repeat this process as necessary. This technique is often called **back-substitution**.

Example 1.5.7. Solve the following linear systems from their augmented matrices which are in echelon form.

(a) Examining below the echelon matrix and its corresponding system of linear equations, we see that we can easily solve this system. Starting with the third equation, we get that $z = 5$. Substituting this into the second equation, which depends only on y and z, we get $y = -46+8z = -46+40 = -6$. Finally, we substitute these values into the first equation and get $x = \frac{1}{2}(y - 3z + 25) = \frac{1}{2}(-6 - 15 + 25) = \frac{4}{2} = 2$. Therefore, the unique solution to the system is $\boxed{(2, -6, 5)}$.

$$\left[\begin{array}{ccc|c} 2 & -1 & 3 & 25 \\ 0 & -1 & 8 & 46 \\ 0 & 0 & 15 & 75 \end{array}\right] \sim \begin{cases} 2x - y + 3z = 25 \\ -y + 8z = 46 \\ 15z = 75 \end{cases}$$

(b) This augmented matrix is in row reduced echelon form. In terms of the linear system, the system is already solved, and the solution corresponds to the augmented column of the matrix, that is, $\boxed{(3, -5, 3)}$ is the unique solution to this linear system.

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 3 \end{array}\right] \sim \begin{cases} x = 3 \\ y = -5 \\ z = 3 \end{cases}$$

(c) The final equation in this system is an identity which is true for all assignments of the variables. It neither adds or takes away any restrictions on the solution set. In fact, this equation could be removed without changing the solution set. Because the system essentially has two equations but three variables, it must be true that at least one of the variables is free.

$$\left[\begin{array}{ccc|c} 1 & 0 & -3 & 2 \\ 0 & 1 & 3 & 7 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \begin{cases} x_1 - 3x_3 = 2 \\ x_2 + 3x_3 = 7 \\ 0 = 0 \end{cases} \sim \begin{cases} x_1 = 2 + 3x_3 \\ x_2 = 7 - 3x_3 \\ x_3 = \text{any number} \end{cases}$$

Notice that the equation $0 = 0$ places no restriction on the variables. It could be removed without altering the solution set. Thus, there is no restriction placed on x_3, that is, x_3 could be freely assigned

any real number t. This is what we mean by calling x_3 a *free variable*. Upon choosing any such number, x_1 and x_2 are determined by their relation with x_3, that is, their assignment is restricted by equations 1 and 2. Thus, this system has multiple solutions, whose general form is $\boxed{(2+3t, 7-3t, t)}$.

Whether there is a free variable or not in the a linear system is relatively easy to determine once you know how to look for it. In particular, dependent variables correspond to pivot columns of the augmented matrix, and free variables correspond to the remaining columns, the so-called **non-pivot columns**.

Theorem 1.5.8. *A linear system has multiple solutions if and only if it is consistent and have at least one non-pivot column. These non-pivot columns will correspond in a one-to-one manner with the free variables of the linear system.*

(d) Examples like the previous one sometimes give the false narrative that multiple solutions occur only when the linear system contains the identity $0 = 0$ in some echelon form. This is patently false. Consider the following *overdetermined* system (more rows than columns) below. The linear system contains two rows of zeros but has a unique solution, namely $(5, -2)$.

$$\left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

On the other hand, an *underdetermined* system (more columns than rows) will necessarily have some columns which cannot have any pivots, as the number of pivots is bounded above by the number of rows and the number of columns. Thus, it MUST have at least one non-pivot column (see Proposition 1.1.8). This translates to mean that an undetermined system MUST have free variables, which will provide multiple solutions if the linear system is consistent. For example, the below linear system is underdetermined but has no row of zeros. Nonetheless, the third column is non-pivot, meaning that x_3 is a free variable of the system. The general solution would then be $(2-t), 3-t, t)$ for any scalar t.

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 2 & 5 \\ 0 & \boxed{1} & 2 & 3 \end{array}\right] \sim \begin{cases} x_1 + x_3 = 2 \\ x_2 + x_3\ 3 \end{cases} \sim \begin{cases} x_1 = 2 - x_3 \\ x_2 = 3 - x_3. \end{cases}$$

The false narrative derivatives from the overexposure linear algebra students often get from $n \times n$ (square) linear systems. In this case, if there is a row of zeros then the square system devolves into an undetermined system which has a free variable. The multiple solutions follow, not from the row of zeros, but from the non-pivot column(s).

(e) The third equation gives a contradiction, since there is no choice of variables such that $0 = 8$. This tells us that the linear system is $\boxed{\text{inconsistent}}$.

$$\left[\begin{array}{ccc|c} 1 & 0 & -1 & 1 \\ 0 & 2 & 4 & 4 \\ 0 & 0 & 0 & 8 \end{array}\right] \sim \begin{cases} x_1 - x_3 = 1 \\ 2x_2 + 4x_3 = 4 \\ 0 = 8 \end{cases}$$

Theorem 1.5.9. *A linear system is inconsistent if and only if it contains a row of the form*

$$\left[\begin{array}{cccc|c} 0 & 0 & \dots & 0 & b \end{array}\right] \qquad (b \neq 0)$$

in some echelon form of the matrix. If the linear system is consistent, then the solution is unique if and only if the system has no free variables. ∎

Exercises

(Go to Solutions)

For Exercises 1-5, each of the augmented matrices are in echelon form. Identify the pivot positions and which variables, e.g. x_1, x_2, $x_3, \ldots$, are free variables in the associated linear system.

1. $\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$

2. $\left[\begin{array}{cccc|c} 1 & 2 & 1 & 4 & -1 \\ 0 & 1 & 3 & 5 & 0 \end{array}\right]$

3. $\left[\begin{array}{cccc|c} 0 & 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 4 & 0 \end{array}\right]$

4. $\left[\begin{array}{cccc|c} 1 & 4 & 2 & 6 & 1 \\ 0 & 2 & 2 & 1 & 2 \\ 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

5. $\left[\begin{array}{ccccc|c} 1 & 3 & 5 & 0 & 1 & 1 \\ 0 & 3 & 4 & 1 & 2 & 2 \\ 0 & 0 & 0 & 6 & 1 & 3 \\ 0 & 0 & 0 & 0 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right]$

For Exercises 6-30, identify if the matrix is in echelon form and if it is in row reduced echelon form. If in echelon form, identify the pivot positions and the rank of the matrix.

6. $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$

7. $\left[\begin{array}{cccc} 1 & 2 & 0 & -3 \\ 0 & 0 & 1 & 5 \end{array}\right]$

8. $\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

9. $\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

10. $\left[\begin{array}{cccc} 3 & -2 & 4 & 5 \end{array}\right]$

11. $\left[\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \end{array}\right]$

12. $\left[\begin{array}{ccc} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$

13. $\left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

14. $\left[\begin{array}{cccc} 1 & 2 & 4 & -3 \\ 0 & 0 & 1 & 5 \end{array}\right]$

15. $\left[\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$

16. $\left[\begin{array}{cccc} 1 & 6 & -5 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$

17. $\left[\begin{array}{cc|c} 5 & 6 & 3 \\ 0 & 4 & 12 \end{array}\right]$

18. $\left[\begin{array}{ccc|c} 2 & 4 & 3 & 6 \\ 0 & 2 & 6 & 2 \\ 0 & 0 & 1 & 4 \end{array}\right]$

19. $\left[\begin{array}{ccc|c} 1 & 0 & 6 & 2 \\ 0 & 0 & 3 & 4 \\ 4 & 0 & 1 & 6 \end{array}\right]$

♠ 20. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 0 & 0 & 3 & 4 \end{array}\right]$

♠ 21. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 1 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$

♠ 22. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 0 & 0 & 2 \\ 0 & 6 & 7 & 8 & 9 \end{array}\right]$

♠ 23. $\left[\begin{array}{cccc|c} 1 & 0 & 3 & 0 & 5 \\ 0 & 1 & 7 & 0 & 9 \\ 0 & 0 & 0 & 1 & 2 \end{array}\right]$

♠ 24. $\left[\begin{array}{cccc|c} 1 & 0 & 3 & 4 & 5 \\ 0 & 1 & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$

♠ 25. $\left[\begin{array}{cccc|c} 1 & 0 & 0 & 4 & 5 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 8 & 9 \end{array}\right]$

26. $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 5 & -1 & 2 & -2 \\ 3 & -1 & 1 & 3 \end{array}\right]$

27. $\left[\begin{array}{ccc|c} 1 & -\frac{1}{5} & \frac{2}{5} & -\frac{2}{5} \\ 0 & 2 & 1 & \frac{7}{3} \\ 0 & 0 & 0 & 1 \end{array}\right]$

28. $\left[\begin{array}{ccc|c} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$

29. $\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 9 & 1 & 3 \\ 0 & 0 & 6 & 7 \\ 2 & 1 & 3 & 4 \end{array}\right]$

30. $\left[\begin{array}{cccc|c} 1 & 0 & 0 & 3 & 5 \\ 0 & 7 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

For Exercises 31-39, write the linear system as augmented matrix or vice versa.

31. $\begin{cases} 4x_2 + 2x_3 = 6 \\ 3x_1 + 5x_2 + 2x_3 = 7 \\ 3x_1 + 17x_2 + 8x_3 = 24 \end{cases}$

♠ 32. $\begin{cases} 3x - y = 0 \\ -2y = 5 \\ x + 7y = 13 \end{cases}$

♠ 33. $\begin{cases} x + y + z = 1 \\ -2x - 6y + z = 12 \end{cases}$

34. $\begin{cases} 10x - 7y + 2z - 4w = 10 \\ 3x + 4y - 3z + w = 20 \\ x + 4z = 73 \\ 6y - 3z + 2w = 10 \end{cases}$

35. $\begin{cases} 12x - 2y + 3z + \frac{3}{2}w = 13 \\ x + 3y - 2z = 9 \\ y + 17z - 20w = 12 \end{cases}$

36. $\left[\begin{array}{ccc|c} 1 & -2 & -3 & -4 \\ 0 & 1 & 2 & 1 \\ -2 & 0 & -2 & 8 \end{array}\right]$

♠ 37. $\left[\begin{array}{ccc|c} 3 & -2 & -1 & 4 \\ 1 & 0 & 3 & -3 \\ 0 & 0 & 0 & 2 \end{array}\right]$

♠ 38. $\left[\begin{array}{cccc|c} 4 & -1 & -2 & 5 & 12 \\ -3 & 0 & 0 & 1 & 5 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right]$

39. $\left[\begin{array}{cccc|c} 2 & 0 & 1 & -9 & 3 \\ 0 & 9 & 21 & 25 & -9 \\ 14 & -7 & 0 & 36 & 67 \\ 3 & 4 & -7 & 19 & 2 \\ -18 & 0 & 2 & 8 & 1 \end{array}\right]$

For Exercises 40-52, perform the indicated elementary row operation(s) to the linear system or augmented matrix.

40. $\begin{cases} 2x + 2y = 5 \\ x + y - z = 1 \\ -5y + 4z = -1 \end{cases}$
scale Row 2 by 3

♠ 41. $\begin{cases} 2y + 5z = 10 \\ x - y + 3z = 6 \\ 2x + 7y - z = 0 \end{cases}$
interchange Rows 1 and 2

♠ 42. $\begin{cases} 3x - y = 0 \\ -2y = 5 \\ x + 7y = 13 \end{cases}$
scale Row 2 by $-1/2$

♠ 43. $\begin{cases} x + y + z = 1 \\ -2x - 6y + z = 12 \end{cases}$
replace Row 2 with
Row 2 + 2Row 1

44. $\begin{cases} 7x + y - 2z \equiv 10 \\ -x - y + 7z \equiv 5 \\ 6x - z \equiv 3 \end{cases} \pmod{11}$
scale Row 2 by $\frac{1}{2}$

45. $\left[\begin{array}{cccc} 1 & 3 & 1 & 3 \\ 0 & 1 & 7 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$
replace Row 1 with
Row 1 + Row 3

46. $\left[\begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 0 \end{array}\right] \left[\begin{array}{c} 0 \\ 5 \\ 7 \end{array}\right]$
replace Row 3 with
Row 3 + 2Row 2

47. $\left[\begin{array}{cccc} 2 & 4 & 6 & -2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} 4 \\ 5 \\ 1 \end{array}\right]$
scale Row 1 by 1/2

48. $\left[\begin{array}{cccc} 2 & 3 & 4 & 6 \\ 1 & 2 & 2 & 3 \\ 9 & 4 & 7 & 2 \end{array}\right]$
replace Row 1 with
Row 1 − Row 2

49. $\left[\begin{array}{cccc} 1 & 2 & 3 & 6 \\ 2 & 3 & 4 & 1 \\ 5 & 6 & 7 & 2 \end{array}\right]$
replace Row 2 with
Row 2 − 2Row 1,
replace Row 3 with
Row 3 − 5Row 1,
replace Row 3 with
Row 3 − 4Row 2

♠ 50. $\left[\begin{array}{cccc} 0 & 0 & 0 & 2 \\ 3 & 2 & 1 & 4 \\ 1 & 0 & 3 & 3 \end{array}\right] \pmod{5}$
interchange Rows 1 and 3

♠ 51. $\left[\begin{array}{cccc} 4 & 4 & 3 & 0 & 2 \\ 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right] \pmod{5}$
scale Row 1 by 4

♠ 52. $\left[\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 2 & 3 \end{array}\right] \left[\begin{array}{c} 0 \\ 2 \\ 4 \end{array}\right] \pmod{5}$
replace Row 1 with Row 1 − 2Row 2

For Exercises 53-57, identify the elementary row operation(s) which transform the first matrix into the second matrix. Answers may vary.

53. $\begin{bmatrix} 0 & 2 & 6 \\ 7 & 1 & 2 \\ -3 & 5 & 0 \end{bmatrix} \sim \begin{bmatrix} 7 & 1 & 2 \\ 0 & 2 & 6 \\ -3 & 5 & 0 \end{bmatrix}$

54. $\begin{bmatrix} 3 & 2 & 0 \\ 5 & 4 & 1 \\ 7 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 \\ 5 & 4 & 1 \\ 3 & 2 & 0 \end{bmatrix}$

55. $\begin{bmatrix} 2 & 0 & 3 \\ 0 & 2 & -2 \\ 3 & 4 & -1 \end{bmatrix} \sim \begin{bmatrix} 1 & 4 & -4 \\ 0 & 1 & -1 \\ 0 & -8 & 11 \end{bmatrix}$

56. $\begin{bmatrix} 1 & -3 & 0 & -9 \\ 3 & -8 & 2 & -21 \\ 2 & -2 & 9 & 7 \\ -2 & 4 & -1 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -3 & 0 & -9 \\ 0 & 1 & 2 & 6 \\ 0 & 4 & 9 & 25 \\ 0 & -2 & -1 & -8 \end{bmatrix}$

57. $\begin{bmatrix} 1 & 0 & 6 & 9 \\ 0 & 1 & 2 & 6 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

[†]The elementary row operations come by many names, and many texts never give them names are all. The replacement operations is sometimes called row addition or row combination because it consists of adding a multiple of a row to another row. The interchange operation is sometimes called row swap or row switch for obvious reasons. Finally, the scaling operation is often called row multiplication because we multiply both sides of an equation by a nonzero constant. The naming scheme used here for the elementary row operations follows the scheme used by Lay [**?**]. We prefer this naming scheme for two reasons. First, students will remember them better and be able to talk with their classmates about them more easily when they have short, simple names. Second, we will eventually see that these three elementary operations correspond to many concepts in linear algebra, including elementary column operations. For this reason, it is preferable to have names which do not use the word "row."

"A leader is someone who helps improve the lives of other people or improve the system they live under."
– Sam Houston

1.6 Reduction of Linear Systems

The key to solving linear systems is to row reduce the augmented matrix to an echelon form. To accomplish this, conduct the following recursive algorithm: identify the leftmost nonzero column. This becomes a pivot column, with the pivot in the top position. If this pivot entry is zero, interchange the pivot row with any nonzero row below it. Next, zero out each entry below the pivot by replacing this row with itself plus a multiple of the pivot row. Consider next the minor matrix where this pivot row and pivot column are removed. Repeat this process on the minor matrix until an echelon form is obtained.

The above algorithm for row reduction to echelon form coupled with the technique of back-substitution to solve the reduced system is call **Gaussian Elimination**. This provides an efficient procedure to solve linear systems.

Example 1.6.1. Reduce the matrix A to echelon form and solve the associated linear system using Gaussian Elimination, where $A = \left[\begin{array}{rrrrr|r} 0 & 0 & 0 & -2 & 3 & -7 \\ -2 & -4 & -6 & 1 & -1 & -3 \\ 4 & 8 & 12 & 1 & -2 & 16 \end{array}\right]$.

We begin by identifying the leftmost nonzero column of the matrix, which is the first column of A. As such, we place a pivot position in the first row of the first column. As this pivot position contains a zero value, we will interchange Rows 1 and 2, since Row 2 contains a nonzero value. (Note that the red double-arrow below will be shorthand we commonly use to denote the interchange of two rows in a matrix.)

$$\left[\begin{array}{rrrrr|r} \boxed{0} & 0 & 0 & -2 & 3 & -7 \\ -2 & -4 & -6 & 1 & -1 & -3 \\ 4 & 8 & 12 & 1 & -2 & 16 \end{array}\right] \sim \left[\begin{array}{rrrrr|r} \boxed{-2} & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & -2 & 3 & -7 \\ 4 & 8 & 12 & 1 & -2 & 16 \end{array}\right]$$

Now we proceed to place zero values in the all the entries below the pivot position in Column 1 using Row Replacement. As Row 2 already has a zero entry in Column 2 we may disregard it for now. To create a zero entry in Row 3 we will replace Row 3 with Row3 $+ 2 \cdot$ Row1, as shown. (Note that red statement "(Row i + cRow j)" written adjacent to Row i will indicate we are replacing Row i with Row i + cRow j. For convenience and to avoid arithmetic errors, we write cRow j write above Row i in red to indicate what values we will add together to produce the next matrix in the row reduction.)

$$\left[\begin{array}{rrrrr|r} \boxed{-2} & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & -2 & 3 & -7 \\ 4 & 8 & 12 & 1 & -2 & 16 \end{array}\right] \begin{array}{l} \\ \\ \text{(Row 3 + 2 Row 1)} \end{array} \sim \left[\begin{array}{rrrrr|r} \boxed{-2} & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & -2 & 3 & -7 \\ 0 & 0 & 0 & 3 & -4 & 10 \end{array}\right]$$

This completes the row reduction in Column 1. We now search for the next pivot position in the minor matrix where Row 1 and Column 1 (the location of the first pivot) are ignored. We notice there are only zero entries remaining in Column 2, so we skip this column. We do the same for Column 3. The leftmost nonzero column is then Column 4. So we place the pivot in the $(2, 4)$ position. As this entry is already

nonzero, no interchange is necessary here.

$$\left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{-2} & 3 & -7 \\ 0 & 0 & 0 & 3 & -4 & 10 \end{array}\right]$$

Next, to remove the 3 below this second pivot we replace Row 3 with Row 3 + $\frac{3}{2}$Row 2. This is demonstrated below.

$$\left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{-2} & 3 & -7 \\ 0 & 0 & 0 & 3 & -4 & 10 \end{array}\right] \sim \left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{-2} & 3 & -7 \\ 0 & 0 & 0 & 0 & 1/2 & -1/2 \end{array}\right]$$

This completes the row reduction in the second pivot column. The third and final pivot would then be placed in the $(3,5)$ position, as this is the leftmost nonzero column when Rows 1 and 2 and Columns 1-4 are ignored. As there are no entries below Row 3, the row reduction is completed automatically.

$$\left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & -2 & 3 & -7 \\ 0 & 0 & 0 & 0 & \boxed{1/2} & -1/2 \end{array}\right]$$

Finally, row reduction ends when we reach the end of the coefficient matrix. Thus, our matrix is in echelon form. In order to solve the associated linear system, we will rewrite the augmented matrix as a system of equations and solve by back-substitution.

$$\begin{cases} -2x_1 - 4x_2 - 6x_3 + x_4 - x_5 = -3 \\ \qquad\qquad\quad - 2x_4 + 3x_5 = -7 \\ \qquad\qquad\qquad\quad \frac{1}{2}x_5 = -\frac{1}{2} \end{cases}$$

■

Remark 1.6.2. Many like to deviate from the Gaussian Elimination algorithm in order to avoid the somewhat more cumbersome computations that follows by having entries which are fractions and having to do arithmetic by hand. One strategy to follow to avoid sometimes unnecessary fractions includes always choosing a row so that the pivot position is a one. Perhaps there is a row with already a one in it. Perhaps a row could be rescaled so there is a row in the pivot position, but if chosen incorrectly we might introduce fractions we are trying to avoid. Finally, a combination of row replacements can produce a pivot one if the greatest common divisor between two entries in the same column is itself one.

For example, in Example 1.6.1 when working on the second pivot, instead of replacing Row 3 with Row 3 + $\frac{3}{2}$Row 2, we could have replaced Row 2 with Row 2 + Row 3 and then replaced Row 3 with Row 3 – 3Row 2, as illustrated below. It takes more row operations but avoiding all fractions.

$$\left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{-2} & 3 & -7 \\ 0 & 0 & 0 & 3 & -4 & 10 \end{array}\right] \sim \left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{1} & -1 & 3 \\ 0 & 0 & 0 & 3 & -4 & 10 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} -2 & -4 & -6 & 1 & -1 & -3 \\ 0 & 0 & 0 & \boxed{1} & -1 & 3 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array}\right]$$

▼

Because the row reduced echelon form is more desirable, we can continue to row reduce a matrix in echelon form that was produced by Gaussian elimination by transforming it into row reduced echelon form. This method is known as **Gauss-Jordan Elimination**. To perform this algorithm, first perform Gaussian elimination to produce an echelon form. This is known as the *forward phase*. For the *backward phase*, rescale the rightmost pivot positions to one, if not already done, and create zeros above the pivot position. Repeat this process for each pivot, moving right to left.

Example 1.6.3. Row reduce the matrix A to reduced echelon form, where $A = \left[\begin{array}{cccc|c} 0 & 1 & -3 & -1 & -2 \\ 1 & 1 & 2 & 4 & 3 \\ 3 & 7 & -6 & 8 & 1 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right]$.

We begin with the forward phase of row reduction. We select position $(1,1)$ for the first pivot. As there is a zero in this pivot, we interchange Row 1 and Row 2:

$$\left[\begin{array}{cccc|c} \boxed{0} & 1 & -3 & -1 & -2 \\ 1 & 1 & 2 & 4 & 3 \\ 3 & 7 & -6 & 8 & 1 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right] \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 3 & 7 & -6 & 8 & 1 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right]$$

Next, we replace Row 3 with Row 3 – 3 Row 1:

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 3 & 7 & -6 & 8 & 1 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right] \begin{array}{l} \\ \\ {\scriptstyle -3} \quad {\scriptstyle -3} \quad {\scriptstyle -6} \quad {\scriptstyle -12} \quad {\scriptstyle -9} \\ \\ \end{array} \begin{array}{l} \\ \\ \text{(Row 3 – 3 Row 1)} \\ \\ \end{array} \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 0 & 4 & -12 & -4 & -8 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right]$$

Next, we move the pivot to position $(2,2)$. Using this pivot, we replace Row 3 with Row 3 – 4 Row 2 and replace Row 4 with Row 4 + Row 1:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & \boxed{1} & -3 & -1 & -2 \\ 0 & 4 & -12 & -4 & -8 \\ 0 & -1 & 3 & 4 & -4 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & \boxed{1} & -3 & -1 & -2 \\ 0 & {\scriptstyle -4} & -12 & {\scriptstyle 4} & {\scriptstyle 8} \\ 0 & {\scriptstyle 1} & 3 & {\scriptstyle -1} & {\scriptstyle -2} \end{array}\right] \begin{array}{l} \\ \\ \text{(Row 3 – 4 Row 1)} \\ \text{(Row 4 + Row 2)} \end{array} \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & \boxed{1} & -3 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & -6 \end{array}\right]$$

We next move the pivot to position $(3,4)$ as there are only zeros in $(3,3)$ and $(4,3)$. As the third row is a zero row, we next interchange Rows 3 and 4. We note that this matrix is now in echelon form. To continue to RREF, we scale Row 3 by $\frac{1}{3}$.

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 0 & 0 & 0 & \boxed{0} & 0 \\ 0 & 0 & 0 & 3 & -6 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 0 & 0 & 0 & \boxed{3} & -6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} \\ \\ \frac{1}{3}\text{(Row 3)} \\ \\ \end{array} \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 0 & 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We begin now the backward phase of row reduction. In order to zero-out the entries above the pivot, we replace Row 1 with Row 1 – 4 Row 3 and Row 2 with Row 2 + Row 3:

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 4 & 3 \\ 0 & 1 & -3 & -1 & -2 \\ 0 & 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} {\scriptstyle -4} \\ {\scriptstyle 1} \\ \\ \\ {\scriptstyle 8} \\ {\scriptstyle -2} \\ \\ \end{array} \begin{array}{l} \\ \text{(Row 1 – 4 Row 3)} \\ \text{(Row 2 + Row 3)} \\ \\ \end{array} \sim \left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 11 \\ 0 & 1 & -3 & 0 & -4 \\ 0 & 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

We next move back to the pivot in position $(2,2)$. As this entry is already one, we proceed to eliminate the one above it. This is accomplished by replacing Row 1 with Row 1 – Row 2. This final matrix is the row reduced echelon form of the original matrix. All the pivot columns are indicated.

$$\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 11 \\ 0 & \boxed{1} & -3 & 0 & -4 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{l} {\scriptstyle -1} \\ \\ \\ \\ {\scriptstyle 3} \\ \\ \\ \\ {\scriptstyle 4} \\ \\ \\ \\ \end{array} \begin{array}{l} \\ \text{(Row 1 – Row 2)} \\ \\ \\ \end{array} \sim \left[\begin{array}{cccc|c} \boxed{1} & 0 & 5 & 0 & 15 \\ 0 & \boxed{1} & -3 & 0 & -4 \\ 0 & 0 & 0 & \boxed{1} & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

■

Theorem 1.6.4. *Each matrix is row equivalent to one and only one reduced echelon matrix.*

The proof of this important theorem essentially follows from Gauss-Jordan elimination.

Example 1.6.5. Row-reducing a matrix over an alternative field, such as $\mathbb{C}$ or $\mathbb{Z}_p$ does not change the algebra whatsoever. It only changes the arithmetic associated to this field. To illustrate this fact, we demonstrate how to row reduce a complex matrix.

$$\begin{bmatrix} 4+9i & 6+3i & 0 \\ 1+3i & 2+i & 0 \\ -13+11i & -1+12i & 2+5i \end{bmatrix}$$

Because the first column is nonzero, the first pivot position is in the $(1,1)$-position. If we follow the Gaussian Elimination algorithm, we would replace rows below this pivot by subtracting some multiple of Row 1. But to do this, it will essentially require us to divide Row 1 by $4+9i$. While we certainly can do this, this can be a very cumbersome calculation for humans (especially those who are still a little uncomfortable with complex arithmetic). Because of this, we demonstrate a slight variation to Gaussian Elimination that, while is technically less efficient from a computational point of view, places less cognitive baggage on the reader. Instead of replacing Row 2 with Row $2-\frac{1+3i}{4+9i}$Row 1, we replace Row 1 with Row $1-3$Row2. This will be immediately followed by replacing Rows 2 and 3 with Row $2-(1+3i)$Row 1 and Row $3-(-13+11i)$Row 1, respectively.

$$\begin{bmatrix} \boxed{4+9i} & 6+3i & 0 \\ 1+3i & 2+i & 0 \\ -13+11i & -1+12i & 2+5i \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 1+3i & 2+i & 0 \\ -13+11i & -1+12i & 2+5i \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & 2+i & 0 \\ 0 & -1+12i & 2+5i \end{bmatrix}$$

Our pivot position then moves to the $(2,2)$-position. We again will deviate from the standard algorithm to scale Row 2 by $\frac{1}{2+i}$. From here, replace Row 3 with Row $3-(-1+12i)$Row 2. Finally, with the final

pivot in the $(3,3)$-position, we scale Row 3 by $\frac{1}{2+5i}$, which provides the RREF of the original matrix.

$$\begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{2+i} & 0 \\ 0 & -1+12i & 2+5i \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & -1+12i & 2+5i \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{2+5i} \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-13, compute the row reduced echelon form for the matrix. List the row operations used to transform the matrix to this form. Indicate the first time the matrix is in echelon form. Answers may vary.

1. $\begin{bmatrix} 1 & 2 \\ -4 & 5 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \\ 2 & 4 & 5 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & -2 & -2 \\ 1 & -1 & -4 & -3 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 1 & 5 & 6 & -5 \\ 3 & 0 & -2 & 1 \\ 0 & 1 & 3 & 3 \\ -3 & 5 & 0 & 1 \end{bmatrix}$

5. $\left[\begin{array}{ccc|c} 1 & 2 & 0 & 9 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 5 & 10 \end{array}\right]$

6. $\left[\begin{array}{ccc|c} 8 & 16 & 0 & 14 \\ 0 & 0 & 4 & 7 \\ 0 & 9 & 3 & 9 \end{array}\right]$

7. $\left[\begin{array}{cccc|c} 3 & 0 & 6 & 9 & 18 \\ 0 & 2 & 4 & 0 & 8 \\ 0 & 0 & 0 & 3 & 6 \end{array}\right]$

♠ 8. $\begin{bmatrix} 1+i & -3+7i \\ 2-i & 3+2i \end{bmatrix}$

9. $\begin{bmatrix} 6-2i & 4+12i \\ -5+i & 2-i \end{bmatrix}$

♠ 10. $\begin{bmatrix} 2 & 7 \\ 5 & 11 \end{bmatrix} \pmod{17}$

♠ 11. $\begin{bmatrix} 0 & 5 & 6 \\ 2 & 1 & 2 \\ 6 & 5 & 9 \end{bmatrix} \pmod 7$

12. $\begin{bmatrix} 2 & 5 & 7 \\ 1 & 3 & 6 \\ 0 & 5 & 9 \end{bmatrix} \pmod 7$

13. $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 0 & -1 \\ 3 & -1 & 0 & 4 \\ 1 & 4 & 5 & 1 \end{bmatrix} \pmod 5$

For Exercises 14-18, solve the linear system using Gaussian elimination, that is, convert the linear system into an augmented matrix, reduce it to any echelon form, switch it back into a system of equations, and solve the reduced system using back-substitution.

14. $\begin{cases} x - y + 3z = 1 \\ 2x - 2y + 5z = 6 \\ 2x + 3y - 4z = 2 \end{cases}$

15. $\begin{cases} 2y + 6z = 4 \\ 3y + 2z + x = 5 \\ 3z + 3x + 3y = 3 \end{cases}$

16. $\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ 3x_1 + 8x_2 + 3x_3 = 5 \\ 3x_2 + x_3 = 6 \end{cases}$

17. $\begin{cases} 6x + 3y + 7z \equiv 20 \\ 18y + 9z \equiv 5 \\ 3y + 4z \equiv 10 \\ 3z \equiv 3 \end{cases} \pmod 5$

18. $\begin{cases} x_1 + 3x_2 + 5x_3 + 4x_4 + 6x_5 \equiv 0 \\ 2x_1 + 4x_2 + 4x_3 + 3x_4 + x_5 \equiv 5 \\ x_4 \equiv 5 \\ 6x_1 + 2x_2 + 6x_3 + 3x_4 \equiv 2 \end{cases} \pmod 7$

For Exercises 19-32, solve the linear system using Gauss-Jordan elimination, that is, convert the linear system into an augmented matrix, reduce it to row reduced echelon form, and use this to solve the linear system.

19. $\begin{cases} 2x + 4y \equiv 0 \\ x + 3y \equiv 2 \end{cases} \pmod 5$

20. $\begin{cases} x + 2y = 2 \\ -4y = 8 \\ 2x + 2y = 6 \end{cases}$

♠ 21. $\begin{cases} x + y = 8 \\ x - y = 4 \end{cases}$

♠ 22. $\begin{cases} x - y = 6 \\ 2x - 3z = 16 \\ 2y + z = 4 \end{cases}$

23. $\begin{cases} 3y + 6z = 3 \\ x + 2y - 3z = 2 \\ 4x - 3z = 0 \end{cases}$

24. $\begin{cases} 3x_2 - 2x_3 = 0 \\ x_1 + x_2 + 3x_3 = 12 \\ 6x_1 + 2x_2 + x_3 = 13 \end{cases}$

25. $\begin{cases} x_1 + 4x_2 + 7x_3 = 3 \\ 2x_1 + 5x_2 + 8x_3 = 3 \\ 3x_1 + 6x_2 + 9x_3 = 3 \end{cases}$

♠ 26. $\begin{cases} 2x - 2y - 2z = 2 \\ 2x + 3y + z = 2\,. \\ 3x + 2y = 0 \end{cases}$

27. $\begin{cases} 7x + 8y + 9z = 10 \\ 4x + 5y + 6z = 10 \\ x + 2y + 3z = 10 \end{cases}$

28. $\begin{cases} 2x_1 + x_2 - x_3 = -10 \\ x_1 - 3x_2 - 2x_3 = -4 \\ 3x_1 + x_2 + x_3 = 15 \end{cases}$

29. $\begin{cases} 2x_1 + x_2 + 4x_3 = 199 \\ 2x_1 + 10x_2 + 15x_3 = 984 \\ -x_1 + 10x_3 = 58 \end{cases}$

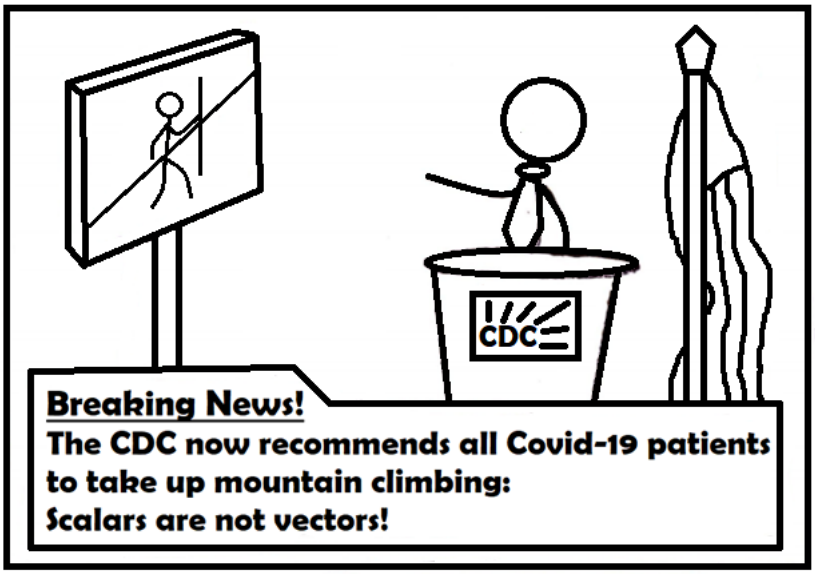
30. $\begin{cases} 5x_1 + 2x_2 + 6x_3 = 11 \\ x_2 + 2x_3 = 5 \\ 3x_1 + 2x_2 + x_3 = 7 \\ 2x_1 + x_2 + 4x_3 = 6 \end{cases}$

31. $\begin{cases} x + y + z + w = 1 \\ 3y + z = 1 \\ x + y + 2z = 0 \\ 2x + z + 4w = 0 \end{cases}$

32. $\begin{cases} 2x = 10 - 4y + 2z \\ -10y = 20 - 16x - 3z - 3 + 13y + 16 - 2z \\ -3z = -4x + 3z - 2y + z \end{cases}$

Chapter 2

Vectors



"If the world's a veil of tears, Smile till rainbows span it." – Lucy Larcom

2.1 Vector Equations

Linear combinations are the bread and butter of vectors. Algebraically speaking, this is what we do with vectors, that is, we combine some vectors to construct new vectors. While it is simple to take a list of vector, combine them, and see what pops out, the converse is not as simple. Given a list of vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in F^m$, can we determine if $\boldsymbol{b} \in F^m$ is a linear combination of these vectors? In other words, we seek to solve the vector equation:

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{b}. \tag{2.1.1}$$

In Equation (2.1.1), note that the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n, \boldsymbol{b}$ are fixed and the variables are the scalars $x_1, x_2, \ldots, x_n \in F$.

To solve the vector equation (2.1.1), suppose that the ith component of $\boldsymbol{a}_j$ is denoted a_{ij}, that is,

$$\boldsymbol{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}.$$

Likewise, let the jth component of $\boldsymbol{b}$ be b_j. Then Equation (2.1.1) becomes

$$\begin{aligned}
x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n &= \boldsymbol{b} \\
x_1\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n\begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} &= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \\
\begin{bmatrix} a_{11}x_1 \\ a_{21}x_1 \\ \vdots \\ a_{m1}x_1 \end{bmatrix} + \begin{bmatrix} a_{12}x_2 \\ a_{22}x_2 \\ \vdots \\ a_{m2}x_2 \end{bmatrix} + \ldots + \begin{bmatrix} a_{1n}x_n \\ a_{2n}x_n \\ \vdots \\ a_{mn}x_n \end{bmatrix} &= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \\
\begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n \end{bmatrix} &= \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
\end{aligned}$$

Examining this last equality of vectors reveals a system of linear equations, namely

$$\begin{cases}
a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\
a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\
\quad\vdots \qquad\quad \vdots \qquad \ddots \qquad \vdots \qquad\quad \vdots \\
a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m,
\end{cases} \tag{2.1.2}$$

whose augmented matrix is

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right].$$

These two problems have the same solution set.

Theorem 2.1.1. *Let $\boldsymbol{x} = (x_1, x_2, \dots, x_m) \in F^m$ be a vector over a field F. Then $\boldsymbol{x}$ is a solution to the vector equation* (2.1.1) *if and only if $\boldsymbol{x}$ is a solution to the linear system* (2.1.2).

Example 2.1.2. Let $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, and $\boldsymbol{a}_3 = \begin{bmatrix} 3 \\ 0 \\ 5 \\ 7 \end{bmatrix}$. Let $\boldsymbol{b} = \begin{bmatrix} -7 \\ 4 \\ -4 \\ -9 \end{bmatrix}$. Is $\boldsymbol{b}$ a linear combination of $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$?

To solve the vector equation,

$$x_1 \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 0 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 4 \\ -4 \\ -9 \end{bmatrix},$$

we solve the linear system whose augmented matrix is

$$\left[\begin{array}{ccc|c} 1 & 1 & 3 & -7 \\ 2 & 0 & 0 & 4 \\ 3 & 0 & 5 & -4 \\ 4 & 1 & 7 & -9 \end{array}\right].$$

We next row reduce the matrix via the elementary row operations:

$$\left[\begin{array}{ccc|c} \boxed{1} & 1 & 3 & -7 \\ 2 & 0 & 0 & 4 \\ 3 & 0 & 5 & -4 \\ 4 & 1 & 7 & -9 \end{array}\right] \begin{array}{l} \\ (\text{Row } 2 - 2\text{Row } 1) \\ (\text{Row } 3 - 3\text{Row } 1) \\ (\text{Row } 4 - 4\text{Row } 1) \end{array} \sim \left[\begin{array}{ccc|c} \boxed{1} & 1 & 3 & -7 \\ 0 & \boxed{-2} & -6 & 18 \\ 0 & -3 & -4 & 17 \\ 0 & -3 & -5 & 19 \end{array}\right] \begin{array}{l} \\ (-\frac{1}{2}\text{Row } 2) \\ \\ (\text{Row } 4 - \text{Row } 3) \end{array}$$

$$\sim \left[\begin{array}{ccc|c} \boxed{1} & 1 & 3 & -7 \\ 0 & \boxed{1} & 3 & -9 \\ 0 & -3 & -4 & 17 \\ 0 & 0 & -1 & 2 \end{array}\right] \begin{array}{l} \\ \\ (\text{Row } 3 + 3\text{Row } 2) \\ (-\text{Row } 4) \end{array} \sim \left[\begin{array}{ccc|c} \boxed{1} & 1 & 3 & -7 \\ 0 & \boxed{1} & 3 & -9 \\ 0 & 0 & \boxed{5} & -10 \\ 0 & 0 & 1 & -2 \end{array}\right] \begin{array}{l} \\ \\ \\ (\text{Interchange}) \\ (\text{Interchange}) \end{array}$$

$$\sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 3 & & -7 \\ 0 & \boxed{1} & 3 & & -9 \\ 0 & 0 & \boxed{1} & & -2 \\ 0 & 0 & 5 & & -10 \end{array}\right] \quad (\text{Row } 4 - 5\text{Row } 3) \qquad \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 3 & & -7 \\ 0 & \boxed{1} & 3 & & -9 \\ 0 & 0 & \boxed{1} & & -2 \\ 0 & 0 & 0 & & 0 \end{array}\right]$$

This last matrix is in echelon form. We will continue to row reduced echelon form.

$$\sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 3 & & -7 \\ 0 & \boxed{1} & 3 & & -9 \\ 0 & 0 & \boxed{1} & & -2 \\ 0 & 0 & 0 & & 0 \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \end{array} \begin{array}{l} (\text{Row } 1 - 3\text{Row } 3) \\ (\text{Row } 2 - 3\text{Row } 3) \\ \\ \end{array} \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 0 & & -1 \\ 0 & \boxed{1} & 0 & & -3 \\ 0 & 0 & \boxed{1} & & -2 \\ 0 & 0 & 0 & & 0 \end{array}\right] \quad (\text{Row } 1 - \text{Row } 2) \qquad \sim \left[\begin{array}{cccc|c} \boxed{1} & 0 & 0 & & 2 \\ 0 & \boxed{1} & 0 & & -3 \\ 0 & 0 & \boxed{1} & & -2 \\ 0 & 0 & 0 & & 0 \end{array}\right]$$

Therefore, the solution is $(2, -3, -2)$. In fact, we see that

$$2\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - 3\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} - 2\begin{bmatrix} 3 \\ 0 \\ 5 \\ 7 \end{bmatrix} = \begin{bmatrix} -7 \\ 4 \\ -4 \\ -9 \end{bmatrix}. \qquad \blacksquare$$

Example 2.1.3. Let $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$, $\boldsymbol{a}_3 = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$, $\boldsymbol{a}_4 = \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$, and $\boldsymbol{b} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}$ be vectors over $\mathbb{Z}_5^3$. Is $\boldsymbol{b}$ a linear combination of $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, $\boldsymbol{a}_4$?

To solve the vector equation,

$$x_1\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + x_3\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} + x_4\begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix},$$

we work with the augmented matrix

$$\left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 1 & 4 \\ 2 & 0 & 2 & 1 & 2 \\ 3 & 2 & 4 & 4 & 1 \end{array}\right] \begin{array}{l} \\ (\text{Row } 2 - 2\text{Row } 1) \\ (\text{Row } 3 - 3\text{Row } 1) \end{array} \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 1 & 4 \\ 0 & \boxed{3} & 3 & 4 & 4 \\ 0 & 4 & 3 & 1 & 4 \end{array}\right] \quad (\text{Row } 3 - 3\text{Row } 2)$$

$$\sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 1 & 4 \\ 0 & \boxed{3} & 3 & 4 & 4 \\ 0 & 0 & \boxed{4} & 4 & 2 \end{array}\right] \quad (4\text{Row } 3) \qquad \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 2 & 1 & 4 \\ 0 & \boxed{3} & 3 & 4 & 4 \\ 0 & 0 & \boxed{1} & 1 & 3 \end{array}\right] \begin{array}{l} (\text{Row } 1 - 2\text{Row } 3) \\ (\text{Row } 2 - 3\text{Row } 3) \\ \end{array}$$

$$\sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 0 & 4 & 3 \\ 0 & \boxed{3} & 0 & 1 & 0 \\ 0 & 0 & \boxed{1} & 1 & 3 \end{array}\right] \quad (2\text{Row } 2) \quad \sim \left[\begin{array}{cccc|c} \boxed{1} & 1 & 0 & 4 & 3 \\ 0 & \boxed{1} & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & 1 & 3 \end{array}\right] \quad (\text{Row } 1 - \text{Row } 2)$$

$$\sim \left[\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 2 & 3 \\ 0 & \boxed{1} & 0 & 2 & 0 \\ 0 & 0 & \boxed{1} & 1 & 3 \end{array}\right]$$

The resulting linear system is then

$$\begin{cases} x_1 + 2x_4 = 3 \\ x_2 + 2x_4 = 0 \\ x_3 + x_4 = 3 \end{cases} \Rightarrow \begin{cases} x_1 = 3x_4 + 3 \\ x_2 = 3x_4 \\ x_3 = 4x_4 + 3 \end{cases}$$

Therefore, there are five possibilities: $(3,0,3,0)$, $(1,3,2,1)$, $(4,1,1,2)$, $(2,4,0,3)$, and $(0,2,4,4)$. This gives five possible linear combinations such as:

$$3\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}. \qquad \blacksquare$$

Definition 2.1.4. Given vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in F^m$, the **(linear) span** of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$, denoted $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, is the set of all linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ inside of F^m, that is, We say that a subset $S \subseteq F^m$ is **spanned** by $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ if $S = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a **spanning set** for S.

Asking whether a vector $\boldsymbol{b} \in \text{Span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is equivalent to asking whether there exists a solution $\boldsymbol{b}$ to the vector equation

$$x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{b}.$$

(These last two examples need to be replaced/updated).

Example 2.1.5. Let $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$, and $\boldsymbol{b} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$. Is $\boldsymbol{b} \in \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$?

As explained above, answering the question is equivalent to solving the vector equation $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 = \boldsymbol{b}$. Likewise, this vector equation can be solved by row reduction on the augmented matrix $\left[\begin{array}{cc|c} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{b} \end{array}\right]$:

$$\left[\begin{array}{cc|c} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right].$$

Therefore, the solution is $x_1 = 3$ and $x_2 = 2$, that is,

$$\boldsymbol{b} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2 = 3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}. \qquad \blacksquare$$

Example 2.1.6. Let $\boldsymbol{a}_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\boldsymbol{a}_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$, and $\boldsymbol{b} = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$. Is $\boldsymbol{b} \in \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$?

Like above, we row reduce the augmented matrix $\left[\begin{array}{cc|c} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{b} \end{array}\right]$:

$$\left[\begin{array}{cc|c} 1 & 5 & -3 \\ -2 & -13 & 8 \\ 3 & -3 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & -18 & 10 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 5 & -3 \\ 0 & -3 & 2 \\ 0 & 0 & -2 \end{array}\right],$$

which is now in echelon form. The third equation is $0 = -2$, which implies that the system in inconsistent. Therefore, $\boldsymbol{b} \notin \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ ∎

Exercises

(Go to Solutions)

For Exercises 1-4, express the given system of equations as a vector equation.

1. $\begin{cases} x_1 + 3x_2 = 7 \\ 7x_1 + 6x_2 - x_3 = 8 \\ 3x_1 + x_2 + 2x_3 = 3 \end{cases}$

♠ 2. $\begin{cases} 3x_1 + 7x_2 - 5x_3 = -5 \\ -x_1 - 2x_2 + 6x_3 = 4 \end{cases}$

♠ 3. $\begin{cases} -2x_1 + 6x_2 - 11x_3 = 6 \\ x_1 - 2x_2 + 5x_3 = -3 \\ 3x_1 - 19x_2 + 31x_3 = -17 \end{cases}$

4. $\begin{cases} x_1 + 2x_2 + 2x_3 = 1 \\ 3x_1 + x_2 = 4 \\ 5x_1 + 2x_2 + x_3 = 3 \end{cases}$

For Exercises 5-9, determine if $\boldsymbol{b}$ is a linear combination of the other vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}$. If so, write $\boldsymbol{b}$ as a linear combination. The set of vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}$ will be listed first, followed by the vector $\boldsymbol{b}$. Answers may vary.

♠ 5. $\left\{ \begin{bmatrix} -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right\}, \begin{bmatrix} 1 \\ 5 \end{bmatrix}$

♠ 6. $\left\{ \begin{bmatrix} -3 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} \right\}, \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$

♠ 7. $\left\{ \begin{bmatrix} 3 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ -3 \end{bmatrix} \right\}, \begin{bmatrix} 1 \\ 6 \\ -9 \end{bmatrix}$

8. $\left\{ \begin{bmatrix} 1+i \\ 1-i \\ 3-2i \end{bmatrix}, \begin{bmatrix} 1+2i \\ 1+i \\ -2+i \end{bmatrix} \right\}, \begin{bmatrix} 4+5i \\ 4-2i \\ 7-5i \end{bmatrix}$

9. $\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 2 \end{bmatrix} \right\}, \begin{bmatrix} 6 \\ 3 \\ 14 \\ 8 \end{bmatrix}$

For Exercises 10-12, determine if $\boldsymbol{b} \in \text{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}$. If so, write $\boldsymbol{b}$ as a linear combination of the $\boldsymbol{a}_i$. The set of vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots\}$ will be listed first, followed by the vector $\boldsymbol{b}$. Answers may vary.

10. $\left\{ \begin{bmatrix} -4 \\ 9 \\ 6 \end{bmatrix}, \begin{bmatrix} 10 \\ 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 7 \\ -12 \\ -8 \end{bmatrix} \right\}, \begin{bmatrix} 12 \\ 4 \\ -6 \end{bmatrix}$

♠ 11. $\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix} \right\}, \begin{bmatrix} -4 \\ -1 \\ -2 \end{bmatrix}$

12. $\left\{ \begin{bmatrix} 4 \\ 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \right\}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \pmod 5$

♠ 13. For any list of vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ in F^m, show that $\text{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ contains the zero vector.

14. Suppose that $\boldsymbol{u}_1, \boldsymbol{u}_2 \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. Prove that $\text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} \subseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$.

15. Suppose that $\boldsymbol{v}_3 \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. Prove that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$.

16. Prove that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2\}$.

17. Suppose that $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Prove that $\text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} \subseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$.
18. Suppose that $\boldsymbol{v}_{n+1} \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Prove that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}\}$.
19. Let $c_i \in F$ be scalars such that $c_n \neq 0$. Prove that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} = \text{Span}\left\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}, \sum_{i=1}^{n} c_i \boldsymbol{v}_n\right\}$.
20. Suppose that $\boldsymbol{u} \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ such that $\boldsymbol{u} \neq \boldsymbol{0}$. Then there exists an i such that

$$\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{u}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}\}.$$

"Some painters transform the sun into a yellow spot, others transform a yellow spot into the sun."
– Pablo Picasso

2.2 Matrix Equations

Definition 2.2.1. Let A be an $m \times n$ matrix and let $\boldsymbol{x} \in F^n$. Let the **column vectors** of A be $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n \in F^n$, that is, $A = \left[\begin{array}{cccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{array}\right]$. Then the **product** $A\boldsymbol{x}$ is the linear combination of the column vectors of A with coefficients corresponding to the entries of $\boldsymbol{x}$, that is,

$$A\boldsymbol{x} = \left[\begin{array}{cccc} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{array}\right] \left[\begin{array}{c} x_1 \\ x_2 \\ \vdots \\ x_n \end{array}\right] = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n. \tag{2.2.1}$$

Note that $A\boldsymbol{x}$ is only defined if the number of columns of A is equal to the number of entries of $\boldsymbol{x}$.

Example 2.2.2.

(a) $$\left[\begin{array}{ccc} 1 & 2 & -3 \\ 0 & -2 & 5 \end{array}\right] \left[\begin{array}{c} 3 \\ 0 \\ 2 \end{array}\right] = 3\left[\begin{array}{c} 1 \\ 0 \end{array}\right] + 0\left[\begin{array}{c} 2 \\ -2 \end{array}\right] + 2\left[\begin{array}{c} -3 \\ 5 \end{array}\right] = \left[\begin{array}{c} 3 \\ 0 \end{array}\right] + \left[\begin{array}{c} 0 \\ 0 \end{array}\right] + \left[\begin{array}{c} -6 \\ 10 \end{array}\right] = \left[\begin{array}{c} -3 \\ 10 \end{array}\right].$$

(b) $$\left[\begin{array}{cc} 2 & 0 \\ -6 & 1 \\ 3 & 2 \end{array}\right] \left[\begin{array}{c} 3 \\ 5 \end{array}\right] = 3\left[\begin{array}{c} 2 \\ -6 \\ 3 \end{array}\right] + 5\left[\begin{array}{c} 0 \\ 1 \\ 2 \end{array}\right] = \left[\begin{array}{c} 6 \\ -18 \\ 9 \end{array}\right] + \left[\begin{array}{c} 0 \\ 5 \\ 10 \end{array}\right] = \left[\begin{array}{c} 6 \\ -13 \\ 19 \end{array}\right].$$

■

Theorem 2.2.3. *If A is an $m \times n$ matrix, with column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$, and if $\boldsymbol{b} \in F^m$, the matrix equation*

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.2.2}$$

has the same solution set as the vector equation

$$x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{b}$$

which, in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$\left[\begin{array}{ccc|c} \boldsymbol{a}_1 & \ldots & \boldsymbol{a}_n & \boldsymbol{b} \end{array}\right].$$

Example 2.2.4. The solution set to the system of equation

$$\begin{cases} x_1 + 3x_2 - 7x_3 = 5 \\ \quad - 2x_2 + 11x_3 = 3 \end{cases}$$

is same as the solution set of the vector equation

$$x_1\begin{bmatrix}1\\0\end{bmatrix}+x_2\begin{bmatrix}3\\-2\end{bmatrix}+x_3\begin{bmatrix}-7\\11\end{bmatrix}=\begin{bmatrix}5\\3\end{bmatrix},$$

which in turn has the same solution set as the matrix equation

$$\begin{bmatrix}1&3&-7\\0&-2&11\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}5\\3\end{bmatrix}.$$ ■

Corollary 2.2.5. *The equation $A\boldsymbol{x}=\boldsymbol{b}$ has a solution if and only if $\boldsymbol{b}$ is a linear combination of the column vectors of A.*

Definition 2.2.6. Let $\mathrm{Col}(A)$ denote the set of all linear combinations of column vectors of A, which is called the **column space** of A. In particular, if $A=\begin{bmatrix}\boldsymbol{a}_1&\boldsymbol{a}_2&\dots&\boldsymbol{a}_n\end{bmatrix}$, then $\mathrm{Col}(A)=\mathrm{Span}\{\boldsymbol{a}_1,\boldsymbol{a}_2,\dots,\boldsymbol{a}_n\}$.

According to Corollary 2.2.5, the equation $A\boldsymbol{x}=\boldsymbol{b}$ is consistent if and only if $\boldsymbol{b}\in\mathrm{Col}(A)$.

Example 2.2.7. Let $A=\begin{bmatrix}1&3&1\\2&4&4\\3&5&7\end{bmatrix}$. Compute the column space of A.

The quick response to this example would be to find a spanning set for the column space of A, but, by definition, a spanning set for a column space is never mysterious. Note that

$$\mathrm{Col}(A)=\mathrm{Span}\left\{\begin{bmatrix}1\\2\\3\end{bmatrix},\begin{bmatrix}3\\4\\5\end{bmatrix},\begin{bmatrix}1\\4\\7\end{bmatrix}\right\}.$$

While we have technically computed the column space, that is, we have found a description of the set of vectors, it does not leave the decision question of whether a generic vector $\boldsymbol{b}=(b_1,b_2,b_3)$ belongs to $\mathrm{Col}(A)$ well answered, which is really what we want to be able to do. To answer this question, we row reduce the augmented matrix

$$\left[\begin{array}{ccc|c}1&3&1&b_1\\2&4&4&b_2\\3&5&7&b_3\end{array}\right]\sim\left[\begin{array}{ccc|c}1&3&1&b_1\\0&-2&2&b_2-2b_1\\0&-4&4&b_3-3b_1\end{array}\right]\sim\left[\begin{array}{ccc|c}1&3&1&b_1\\0&1&1&b_1-\frac{1}{2}b_2\\0&1&-1&\frac{1}{4}(3b_1-b_3)\end{array}\right]\sim\left[\begin{array}{ccc|c}1&3&1&b_1\\0&1&1&b_1-\frac{1}{2}b_2\\0&0&0&-\frac{1}{4}b_1+\frac{1}{2}b_2-\frac{1}{4}b_3\end{array}\right]$$

Therefore, the above equation is consistent if and only if $-\frac{1}{4}b_1+\frac{1}{2}b_2-\frac{1}{4}b_3=0$. Therefore, Now, there exists plenty of choices of $\boldsymbol{b}$ such that this equation does not hold. Hence, $A\boldsymbol{x}=\boldsymbol{b}$ is not consistent for all $\boldsymbol{b}$.

$$\mathrm{Col}(A)=\{\boldsymbol{b}\in\mathbb{R}^3\mid b_1-2b_2+b_3=0\}.$$ ■

Generalizing the principles seen in the previous example, for any $n \times n$ matrix A, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a solution for all $\boldsymbol{b}$ if and only if A has a pivot in each row, in which case the solution must be unique.

Let A be an $(m \times n)$ matrix and let $\boldsymbol{x} \in F^n$. Then the rule $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is a transformation called a *matrix transformation.*

Theorem 2.2.8. *Let A be an $m \times n$ matrix and let $\boldsymbol{u}, \boldsymbol{v} \in F^n$. Let $c \in F$. Then $A(\boldsymbol{u} + \boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$ and $A(c\boldsymbol{u}) = c(A\boldsymbol{u})$. In particular, multiplication by a matrix is a linear transformation.*

Let A be an $m \times n$ matrix with column vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$, that is, let a_{ij} denote the entry of A in the (i, j) position. Let $\boldsymbol{x} \in F^n$. Then

$$A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + \ldots + x_n\boldsymbol{a}_n = x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \ldots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= \begin{bmatrix} x_1a_{11} \\ x_1a_{21} \\ \vdots \\ x_1a_{m1} \end{bmatrix} + \begin{bmatrix} x_2a_{12} \\ x_2a_{22} \\ \vdots \\ x_2a_{m2} \end{bmatrix} + \ldots + \begin{bmatrix} x_na_{1n} \\ x_na_{2n} \\ \vdots \\ x_na_{mn} \end{bmatrix} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \ldots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \ldots + x_na_{mn} \end{bmatrix}.$$

That is

$$A\boldsymbol{x} = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \ldots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \ldots + x_na_{2n} \\ \vdots \\ x_1a_{m1} + x_2a_{m2} + \ldots + x_na_{mn} \end{bmatrix}. \tag{2.2.3}$$

This formula often is easier to compute than the original definition of $A\boldsymbol{x}$.

Example 2.2.9. Let $A = \begin{bmatrix} 0 & -2 \\ 1 & -3 \\ 2 & -3 \end{bmatrix}$ and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\boldsymbol{x}) = A\boldsymbol{x}$. Then

$$T(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} 0 & -2 \\ 1 & -3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

(a) Let $\boldsymbol{u} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$. Compute $T(\boldsymbol{u})$.

$$T(\boldsymbol{u}) = A\boldsymbol{u} = \begin{bmatrix} 0(3) - 2(-1) \\ 1(3) - 3(-1) \\ 2(3) - 3(-1) \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \\ 9 \end{bmatrix}.$$

(b) Find an $\boldsymbol{x} \in \mathbb{R}^2$ whose image under T is $\boldsymbol{b} = \begin{bmatrix} -8 \\ -7 \\ -2 \end{bmatrix}$.

We need an x_1 and x_2 such that

$$\begin{bmatrix} 0 & -2 \\ 1 & -3 \\ 2 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -8 \\ -7 \\ -2 \end{bmatrix}.$$

By row reduction, we solve the corresponding linear system:

$$\left[\begin{array}{cc|c} 0 & -2 & -8 \\ 1 & -3 & -7 \\ 2 & -3 & -2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -3 & -7 \\ 0 & -2 & -8 \\ 2 & -3 & -2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -3 & -7 \\ 0 & -2 & -8 \\ 0 & 3 & 12 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -3 & -7 \\ 0 & 1 & 4 \\ 0 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -3 & -7 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 0 \end{array}\right].$$

Let $\boldsymbol{x} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$. Then, $T(\boldsymbol{x}) = \boldsymbol{b}$. In fact, $\boldsymbol{x}$ is the only vector whose image under T is $\boldsymbol{b}$. ■

Exercises

(Go to Solutions)

For Exercises 1-1, determine with the statement is true or false. If false, correct the statement so that it is true.

1. A homogeneous linear system can be written as $A\boldsymbol{x} = 0$, where A is the $m \times n$ coefficient matrix, $\boldsymbol{x} \in F^n$, and 0 is a scalar.

For Exercises 2-5, compute the matrix-vector product.

2. $\begin{bmatrix} 4 & 5 & 2 \\ 0 & -1 & 3 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 3 & -1 & 2 \\ 4 & 3 & 7 \\ -2 & 1 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 1+2i & -1+3i \\ 0 & 3+4i \end{bmatrix} \begin{bmatrix} 7+i \\ -2-i \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \pmod 2$

For Exercises 6-8, express the given matrix equation as a system of equations.

♠ 6. $\begin{bmatrix} 4 & 1 & 2 & 0 \\ 0 & 4 & 3 & 1 \\ 2 & 2 & 4 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \pmod 5$

♠ 7. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -1 & -2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ -9 \end{bmatrix}$

8. $\begin{bmatrix} 21 & 27 & 17 & -32 \\ 4 & 42 & 6 & 19 \\ 103 & -72 & 17 & -8 \\ 2 & 11 & -10 & 13 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 12 \\ -6 \\ 15 \\ 1 \end{bmatrix}$

For Exercises 9-16, solve the matrix equation.

9. $\begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & -3 \\ -1 & -2 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ -8 \\ -9 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 4 & 4 & 2 \\ -4 & -3 & -2 \\ -4 & -3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 12 \\ -3 \\ 3 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 1 & 0 \\ 2 & 5 \\ 3 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$

12. $\begin{bmatrix} 1-i & 1 & 1+i \\ 1 & i & -i \\ 3+i & 2 & 3i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3i \\ 11+7i \end{bmatrix}$

♠ 13. $\begin{bmatrix} 1 & 4 & 3 \\ 4 & 3 & 4 \\ 3 & 4 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} \pmod 5$

14. $\begin{bmatrix} 0 & 1 & 2 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \pmod 3$

15. $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \pmod 3$

16. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \pmod 2$

For Exercises 17-18, use the transformation $T : \mathbb{R}^4 \to \mathbb{R}^2$ given by the rule:

$$T(x_1, x_2, x_3, x_4) = \begin{bmatrix} 2 & -1 & 0 & 3 \\ 0 & 6 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

♠ 17. Compute $T(1, 0, 0, 1)$ and $T(1, 2, 3, 4)$.

♠ 18. Find $\boldsymbol{x}$ such that $T(\boldsymbol{x}) = (1, 2)$.

19. It can be shown that $(4, 3, 2, 1)$ is the unique solution to the system of linear equations

$$\begin{cases} 11x_1 + 9x_2 + 10x_3 + 14x_4 = 105 \\ 9x_1 + 4x_2 + 5x_3 + 9x_4 = 67 \\ 11x_1 + 12x_2 + 13x_3 + 17x_4 = 123 \\ 7x_1 + 2x_2 + 5x_3 + 7x_4 = 51. \end{cases}$$

With this knowledge, QUICKLY solve the matrix equation

$$\begin{bmatrix} 11 & 9 & 10 & 14 \\ 9 & 4 & 5 & 9 \\ 11 & 12 & 13 & 17 \\ 7 & 2 & 5 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 105 \\ 67 \\ 123 \\ 51 \end{bmatrix}.$$

Why were you able to solve it so quickly?

"All mankind... being all equal and independent, no one ought to harm another in his life, health, liberty or possessions." – John Locke

2.3 Linear Independence

Definition 2.3.1. A set of vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\} \subseteq F^m$ is said to be **linearly independent** if the vector equation

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{0}$$

has only the trivial solution. Otherwise, the set is said to be **linearly dependent**.

In other words, a set of vectors is linearly dependent if there exists some weights $c_1, c_2, \ldots, c_n$, with at least one not zero, such that

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 + \ldots + c_n\boldsymbol{a}_n = \boldsymbol{0}.$$

The previous equation is then called a **linear dependence relation**.

Expressed another way, suppose $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{bmatrix}$, then $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is linearly independent if and only if $A\boldsymbol{x} = \boldsymbol{0}$ has no nontrivial solution, since

$$A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \ldots & \boldsymbol{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \ldots + x_n\boldsymbol{a}_n = \boldsymbol{0}.$$

The set $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\}$ is linearly dependent if and only if $A\boldsymbol{x} = \boldsymbol{0}$ has a nontrivial solution.

Example 2.3.2. Let $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$, and $\boldsymbol{v}_3 = \begin{bmatrix} -3 \\ 4 \\ 13 \end{bmatrix}$.

(a) Determine if the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is linearly independent.

To determine if the set is linearly independent, it suffices to compute the echelon form of the corresponding homogeneous system. In fact, the final row of zeros is irrelevant.

$$\begin{bmatrix} 1 & 3 & -3 \\ 2 & 1 & 4 \\ 3 & -2 & 13 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 \\ 0 & -5 & 10 \\ 0 & -11 & 22 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 \\ 0 & 1 & -2 \\ 0 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus, x_3 is a free variable of the system, which implies that the homogeneous equation $A\boldsymbol{x} = \boldsymbol{0}$ does have a nontrivial solution. Therefore, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is linearly dependent.

(b) Find a linear dependence relation among $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, and $\boldsymbol{v}_3$.

To determine a dependence relation, we finish the row reduction from above.

$$\begin{bmatrix} 1 & 3 & -3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore,

$$\begin{cases} x_1 + 3x_3 = 0 \\ x_2 - 2x_3 = 0 \\ 0 = 0 \end{cases}$$

Thus, $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -3x_3 \\ 2x_3 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$. Thus, $-3\boldsymbol{v}_1 + 2\boldsymbol{v}_2 + \boldsymbol{v}_3 = \boldsymbol{0}$ (corresponding to $x_3 = 1$). Of course, there are infinitely many dependence relations. For example, $-15\boldsymbol{v}_1 + 10\boldsymbol{v}_2 + 5\boldsymbol{v}_3 = \boldsymbol{0}$ (corresponding to $x_3 = 5$).

We also note that we have shown that $\boldsymbol{v}_3 \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ since $-3\boldsymbol{v}_1 + 2\boldsymbol{v}_2 + \boldsymbol{v}_3 = \boldsymbol{0}$ implies $\boldsymbol{v}_3 = 3\boldsymbol{v}_1 - 2\boldsymbol{v}_2$. This observation is always the case for linear dependence. ∎

Example 2.3.3. Determine if the columns of the matrix $A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 2 & 4 \\ 2 & -4 & -7 \end{bmatrix}$ are linearly independent.

It suffices to show that $A\boldsymbol{x} = \boldsymbol{0}$ has no nontrivial solutions, that is, that an echelon form of A has no zero rows.

$$\begin{bmatrix} \boxed{1} & -1 & -1 \\ -1 & 2 & 4 \\ 2 & -4 & -7 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & -1 & -1 \\ 0 & \boxed{1} & 3 \\ 0 & -2 & -5 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & -1 & -1 \\ 0 & \boxed{1} & 3 \\ 0 & 0 & \boxed{1} \end{bmatrix}.$$

Therefore, the columns of A are linearly independent. ∎

These examples illustrate an important point: a set of vector is linearly dependent if and only if the corresponding homogeneous linear system has a free variable. In other words, a set of vectors is linearly independent if and only if the associated coefficient matrix has a pivot in each column.

Theorem 2.3.4.

A set $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$ is linearly dependent if and only if at least one of the vectors of S is a linear combination of the others.

Proof. Suppose first that S is linearly dependent and

$$c_1\boldsymbol{v}_1 + \ldots + c_p\boldsymbol{v}_p = \boldsymbol{0}$$

is a dependence relation. Now there exists some j such that $c_j \neq 0$. Without the loss of generality, we may suppose that $j = p$. Then

$$\begin{aligned} c_1\boldsymbol{v}_1 + \ldots + c_{p-1}\boldsymbol{v}_{p-1} &= -c_p\boldsymbol{v}_p \\ \frac{c_1}{-c_p}\boldsymbol{v}_1 + \ldots + \frac{c_{p-1}}{-c_p}\boldsymbol{v}_{p-1} &= \boldsymbol{v}_p. \end{aligned}$$

So, $\boldsymbol{v}_p \in \text{Span}\{\boldsymbol{v}_1, \ldots \boldsymbol{v}_{p-1}\}$.

Conversely, if $v_p = c_1\boldsymbol{v}_1 + \ldots + c_{p-1}\boldsymbol{v}_{p-1}$, then

$$c_1\boldsymbol{v}_1 + \ldots + c_{p-1}\boldsymbol{v}_{p-1} - v_p = \boldsymbol{0}$$

is a dependence relation, that is, $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ is linearly dependent. □

The following properties provided, in some cases, a quick test to determine the linear independence or dependence of a set of vectors.

Theorem 2.3.5.

(a) If a set contains a linearly dependent subset, then the set itself is linearly dependent.

(b) The set $\{\boldsymbol{v}\}$ is linearly independent if and only if $\boldsymbol{v} \neq \mathbf{0}$.

(c) If a set $S \subseteq F^n$ contains the zero vector, then S is linearly dependent.

(d) The set $\{\boldsymbol{u}, \boldsymbol{v}\}$ is linearly independent if and only if neither vector is a multiple of the other, that is $\boldsymbol{v} \neq c\boldsymbol{u}$ nor $\boldsymbol{u} \neq c\boldsymbol{v}$ for any $c \in F$.

(e) Suppose $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$. If $p > n$, then S is linearly dependent.

Proof.

(a) Suppose S is linearly dependent and $S \subseteq T$. Then set every vector in T which is not in S to have weight zero. Then a dependence relation on S becomes a dependence relation on T.

(b) Note that $c\boldsymbol{v} = \mathbf{0}$ if and only if $c = 0$ or $\boldsymbol{v} = \mathbf{0}$.

(c) Set the weight of the zero vector to one and set the weight of any other vector in S to zero. This gives a dependence relation.

(d) This follows immediately from Theorem 2.3.4 when $p = 2$.

(e) Let $A = \begin{bmatrix} \boldsymbol{v}_1 & \cdots & \boldsymbol{v}_p \end{bmatrix}$. Then A is an $n \times p$ matrix and the equation $A\boldsymbol{x} = \mathbf{0}$ corresponds to an under-determined system of n equation with p unknowns. Thus, the system has a free variable by Theorem 1.1.8, which implies the linear dependence of the column vectors. □

Example 2.3.6. Determine by inspection if the given sets are linearly independent.

(a) $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\}$.

This set of vectors is actually just a single vector. As such, the set is linearly independent exactly when this vector is nonzero, which it is. Therefore, the set is linearly independent.

(b) $\left\{ \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 7 \end{bmatrix} \right\}$.

This set of vectors is linearly dependent because it contains the zero vector. Note that $0\begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 2 \\ 1 \\ 7 \end{bmatrix} = \mathbf{0}$.

(c) $\left\{\begin{bmatrix} 15 \\ 0 \\ 20 \\ -5 \end{bmatrix}, \begin{bmatrix} -12 \\ 0 \\ -16 \\ 4 \end{bmatrix}\right\}$.

This set of vectors is linearly dependent because the second vector is just $(-5/4)$ times the first vector.

(d) $\left\{\begin{bmatrix} 7 \\ 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 9 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 9 \end{bmatrix}\right\}$.

This set of vectors is linearly dependent because any set of 4 vectors in $\mathbb{R}^3$ must be dependent.

■

Exercises

(Go to Solutions)

For Exercises 1-6, determine with the statement is true or false. If false, correct the statement so that it is true.

1. If a set $S \subseteq F^n$ contains the zero vector, then S is linearly independent.
2. If a set contains a linearly dependent subset, then the set itself is linearly dependent.
3. The set $\{\boldsymbol{u}, \boldsymbol{v}\}$ is linearly independent if and only if neither vector is a multiple of the other, that is, $\boldsymbol{v} \neq c\boldsymbol{u}$ and $\boldsymbol{u} \neq c\boldsymbol{v}$ for any $c \in F$.
4. Suppose $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$. If $p > n$, then S is linearly dependent.
5. The set $\{\boldsymbol{v}\}$ is linearly independent if and only if $\boldsymbol{v} \neq 0$.
6. A set S of two or more vectors is linearly dependent if and only if at least one of the vectors in S is a linear combination of the others.

QUICK! For Exercises 7-13, determine whether the set of vectors is linearly independent or dependent using Theorem 2.3.5 in LESS THAN 10 SECONDS! Justify your quick response.

♠ 7. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$

♠ 8. $\left\{ \begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$

9. $\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 6 \\ 7 \end{bmatrix}, \begin{bmatrix} 7 \\ 1 \\ 0 \\ 2 \end{bmatrix} \right\}$

10. $\left\{ \begin{bmatrix} -1+i \\ 2 \end{bmatrix}, \begin{bmatrix} -6+3i \\ 3+i \end{bmatrix}, \begin{bmatrix} i \\ -4+2i \end{bmatrix} \right\}$

♠ 11. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\} \pmod{11}$

♠ 12. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} \pmod{2}$

13. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 2 \\ 2 \end{bmatrix} \right\} \pmod{5}$

For Exercises 14-20, determine if the set of vectors is linearly independent or dependent. If linearly dependent, provide a nontrivial dependency relation.

14. $\left\{ \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 5 \end{bmatrix} \right\}$

♠ 15. $\left\{ \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 6 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix} \right\}$

♠ 16. $\left\{\begin{bmatrix} 1 \\ -2 \\ -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ -3 \\ 2 \end{bmatrix}\right\}$

17. $\left\{\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}\right\}$

♠ 18. $\left\{\begin{bmatrix} i \\ 1 \\ -i \end{bmatrix}, \begin{bmatrix} 2 \\ 1-2i \\ -1+i \end{bmatrix}, \begin{bmatrix} 0 \\ i \\ i \end{bmatrix}\right\}$

♠ 19. $\left\{\begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\right\} \pmod 5$

♠ 20. $\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 6 \\ 6 \\ 4 \end{bmatrix}, \begin{bmatrix} 6 \\ 5 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 5 \\ 4 \\ 4 \end{bmatrix}\right\} \pmod 7$

"Space is an inspirational concept that allows you to dream big." – Peter Diamandis

2.4 Affine Geometry

For any field, there is a limited amount of geometry, called **affine geometry**, we can attach to the vector space F^n by mimicking geometric structures from $\mathbb{R}^n$.

Definition 2.4.1. Let F^n be a vector space. Then a **flat** (or **affine set**) is a subset of F^n which is congruent to F^m for some $0 \leq m \leq n$. Equivalently, flats are solution sets of linear systems (vector equations, matrix equations, etc.).

To describe flats, e.g., points, lines, planes, we present two recursive constructions: *Top-Down* or *Bottom-Up*. The *Top-Down* approach is to start with a single non-zero, linear equation $a_{1,1}x_1 + \ldots + a_{1,n}x_n = b_i$. The solution set of this linear equation forms a special kind of flat, called a **hyperplane**. For example, $ax + by + cz = d$ defines a *plane* in $\mathbb{R}^3$. By plane, we mean something that "looks" like $\mathbb{R}^2$ inside of $\mathbb{R}^n$. In general, a *plane* over F^n should be a subset that "looks" like F^2. The solution set to this $1 \times n$ system should have $n-1$ free variables. Generally speaking, we view a hyperplane in F^n as an affine set that "looks" likes F^{n-1} in F^n.

Next, consider a linear system that contains the above linear equation and a second one $a_{2,1}x_1 + \ldots + a_{2,n}x_n = b_2$. The solution set to this $2 \times n$ linear system is geometrically the intersection of the two hyperplanes, for example, two distinct planes intersecting in $\mathbb{R}^3$ form a *line*. By line, we mean something that looks like $\mathbb{R}^1 = \mathbb{R}$ inside of $\mathbb{R}^n$. In general, a *line* over F^n should be a subset that "looks" like $F^1 = F$. If the equations are linearly independent, then this $2 \times n$ system will have $n-2$ free variables. Hence, we view an intersection of two hyperplanes in F^n as an affine set that "looks" likes F^{n-2} in F^n.

Continuing in this fashion of expanding the linear system by adding new linear equations while maintaining linear independence, the $m \times n$ linear system will have $n-m$ free variables and the flat will resemble F^{n-m}. This continues until $m = n$ and the intersection of hyperplanes is just a point, which resembles $F^0 = \{\mathbf{0}\}$. In this case, there are no free variables to the linear system. This is, of course, only true if the set of linear equations is linearly independent; if linearly dependent, then at least one equation is a linear combination of the rest and its removal does not change the flat one bit. That is, the inclusion of a linear combination of equations already in the system offers no restriction on the solution set whatsoever, and this equation is frankly redundant to the linear system. This summarizes the *Top-Down* approach to affine sets.

The "shape" of the flat, that is, the parameter p for which the affine set resembles F^p, appears to be determined by the number of free variables in the linear system. The *Bottom-Up* approach to flats is to adjoin more and more free variables until the flat has the right "shape."

A **point** in F^n is the solution to the vector equation

$$\boldsymbol{x} = \boldsymbol{x}_0,$$

where $\boldsymbol{x}$ is a variable vector and $\boldsymbol{x}_0$ is a fixed vector. So a point in F^n is just a vector, that is, $P = \boldsymbol{x}_0$. It is also a translation of the empty span $\text{Span}\{\} = \{\mathbf{0}\} = F^0$.

A **line** in F^n, call it ℓ, is the translation of the span of a single vector $\boldsymbol{v}$, which acts as the direction or slope of the line. If $\boldsymbol{x}_0$ is a point on ℓ, then ℓ is the solution set to the vector equation

$$\boldsymbol{x} = \boldsymbol{x_0} + t\boldsymbol{v}.$$

Similarly, a **plane** in F^n, call it $\mathcal{P}$, is a translation of a span of two linearly independent vectors, that is, a translation of $\text{Span}\{\boldsymbol{u}, \boldsymbol{v}\}$. Then $\mathcal{P}$ is the solution set to the vector equation

$$\boldsymbol{x} = \boldsymbol{x_0} + s\boldsymbol{u} + t\boldsymbol{v},$$

where vectors $\boldsymbol{u},\ \boldsymbol{v} \in F^n$, $\{\boldsymbol{u}, \boldsymbol{v}\}$ is linear independent, and $\boldsymbol{x_0}$ is on $\mathcal{P}$.

By allegory, we can construct higher flats by increasing the number of linearly independent vectors in the linear combination associated with the vector equation, that is, we increase the number of linearly independent vectors in the spanning set, the so-called **spanners** of the flat. For example, a hyperplane in F^4 is the set of solutions to the vector equation

$$\boldsymbol{x} = \boldsymbol{x}_0 + r\boldsymbol{u} + s\boldsymbol{v} + t\boldsymbol{w},$$

where $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\} \subseteq F^4$ is linearly independent. More generally, an m-flat is the solution set to the vector equation

$$\boldsymbol{x} = \boldsymbol{x}_0 + \sum_{i=1}^{m} t_i \boldsymbol{v}_i, \tag{2.4.1}$$

where $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\} \subseteq F^n$ is a linearly independent set of vectors. Equation (2.4.1) is called the **vector form of the flat**. Each component in this vector equation is a linear equation in its own right. The system of linear equations associated to this vector equation is called the **parametric equations** of the flat. These parametric equations would be the general solution to the $(n-1) \times n$ *Top-Down* linear system described above. This summarizes the *Bottom-Up* approach to affine sets.

Example 2.4.2. Find a vector equation and parametric equations of the line in $\mathbb{R}^3$ that passes through the point $\boldsymbol{x_0} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and is parallel to the vector $\boldsymbol{v} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}$.

If the flat is said to be "parallel" to a vector, then that vector acts as a spanner for the flat. Hence, the vector equation is simple enough:

$$\boldsymbol{x} = \boldsymbol{x_0} + t\boldsymbol{v}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}.$$

For the parametric equations, we look closer:

$$\begin{cases} x_1 &= 1 + 5t \\ x_2 &= 2 - 3t \\ x_3 &= 3 + \ \ t. \end{cases}$$

■

Example 2.4.3. Find a vector equation and parametric equations of the line in $\mathbb{R}^4$ that passes through the origin and is parallel to the vector $\boldsymbol{v} = (4, -3, 2, -1)$.

As the line passes through the origin, we set $\boldsymbol{x}_0 = \boldsymbol{0}$. Hence, the vector equation for this line is :

$$\boldsymbol{x} = t\boldsymbol{v},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} 4 \\ -3 \\ 2 \\ -1 \end{bmatrix},$$

and the the parametric equations are given as:

$$\begin{cases} x_1 &= \ \ 4t \\ x_2 &= -3t \\ x_3 &= \ \ 2t \\ x_4 &= \ -t. \end{cases}$$

■

Example 2.4.4. Find a vector equation and parametric equations of the plane in $\mathbb{R}^4$ that passes through $\boldsymbol{x}_0 = (26, 3, -13, -18)$ and is parallel to both the vectors $\boldsymbol{u} = (1, -3, -2, -1)$ and $\boldsymbol{v} = (0, 0, 1, 0)$.

The vector equation takes on the form:

$$\boldsymbol{x} = \boldsymbol{x_0} + s\boldsymbol{u} + t\boldsymbol{v},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 26 \\ 3 \\ -13 \\ -18 \end{bmatrix} + s \begin{bmatrix} 1 \\ -3 \\ -2 \\ -1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

and the the parametric equations are given as:

$$\begin{cases} x_1 &= 26 + s \\ x_2 &= 3 - 3s \\ x_3 &= -13 - 2s + t \\ x_4 &= -18 - s \end{cases}.$$

■

Suppose we have a line which contains the vector $\boldsymbol{x}_0$ and with spanner $\boldsymbol{v}$. Let $\boldsymbol{x}_1$ be another vector on this line. Then there exists some $s \in F$ such that $\boldsymbol{x}_1 = \boldsymbol{x}_0 + s\boldsymbol{v}$. Note that $s \neq 0$, since $\boldsymbol{x}_1 \neq \boldsymbol{x}_0$. Hence, $s\boldsymbol{v} = \boldsymbol{x}_1 - \boldsymbol{x}_0$, which implies that the difference of any two vectors on the line is the a scalar multiple of the spanner. Also, $\boldsymbol{v} = \frac{1}{s}(\boldsymbol{x}_1 - \boldsymbol{x}_0)$. Thus, for any vector $\boldsymbol{x}$ on the line, we have

$$\boldsymbol{x} = \boldsymbol{x}_0 + t\boldsymbol{v} = \boldsymbol{x}_0 + \frac{t}{s}(\boldsymbol{x}_1 - \boldsymbol{x}_0) = \frac{s-t}{s}\boldsymbol{x}_0 + \frac{t}{s}\boldsymbol{x}_1.$$

Note that the coefficients satisfy the condition that $\frac{s-t}{s} + \frac{t}{s} = 1$. In general, the line containing $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ is the set of vectors of the form $\{a_0\boldsymbol{x}_0 + a_1\boldsymbol{x}_1 \mid a_0, a_1 \in F, a_0 + a_1 = 1\}$.

These principles can be generalized to higher flats too.

Definition 2.4.5. Given vectors $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in F^n$, the vector $\boldsymbol{x}$ given as

$$\boldsymbol{x} = a_0\boldsymbol{x}_0 + \boldsymbol{a}_1\boldsymbol{x}_1 + \ldots + a_m\boldsymbol{x}_m$$

is called an **affine combination** of $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ if the scalars a_i satisfy the equality $a_0 + a_1 + \ldots + a_m = 1$. The **affine span** of $\{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$, denoted $\mathrm{Aff}(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m)$, is the set of all affine combinations of $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, that is,

$$\mathrm{Aff}(\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m) = \left\{ \sum_{i=0}^{m} a_i\boldsymbol{x}_i \;\middle|\; \sum_{i=0}^{m} a_i = 1 \right\}.$$

Note that the affine span of the vectors $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ is not the same object as the linear span of the vectors $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ (recall Definition 2.1.4), although in general

$$\mathrm{Aff}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m\} \subseteq \mathrm{Span}\{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m\}. \tag{2.4.2}$$

Since affine combinations are linear combinations with the extra condition that scalars sum to 1, all affine combinations are linear combinations, but the converse does not hold, hence the above inequality. The affine span of the vectors $\boldsymbol{x}_0, \ldots, \boldsymbol{x}_m$ is the smallest affine set containing these specific vectors. To see this, note that because of the assumption about the coefficients, any affine combination could be written as

$$\boldsymbol{x} = (1 - a_1 - a_2 - \ldots - a_m)\boldsymbol{x}_0 + a_1\boldsymbol{x}_1 + a_2\boldsymbol{x}_2 + \ldots + a_m\boldsymbol{x}_m.$$

On the other hand, the affine combination $\boldsymbol{x}$ is on the flat determined by the equation

$$\boldsymbol{x} = \boldsymbol{x}_0 + a_1(\boldsymbol{x}_1 - \boldsymbol{x}_0) + a_2(\boldsymbol{x}_2 - \boldsymbol{x}_0) + \ldots + a_m(\boldsymbol{x}_m - \boldsymbol{x}_0). \tag{2.4.3}$$

Notice this flat also contains each $\boldsymbol{x}_j$, which is obtained by setting $a_i = 1$ when $i = j$ and $a_i = 0$ when $i \neq j$.

We see the very important observation from (2.4.3): the spanners for a flat can be found by taking differences of specific vectors on the flat.

Example 2.4.6. Find a vector equation and parametric equations of the line in $\mathbb{R}^3$ through $(1, 2, 3)$ and $(2, -2, 0)$.

We can take $\boldsymbol{x}_0$ to be either of the two points. We will take $\boldsymbol{x}_0 = (1, 2, 3)$. To get the spanner $\boldsymbol{v}$, we need to take the difference of the two vectors provided, that is, $\boldsymbol{v} = (2, -2, 0) - (1, 2, 3) = (1, -4, -3)$.

Therefore, the vector equation is given as

$$\boldsymbol{x} = \boldsymbol{x_0} + t\boldsymbol{v},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ -4 \\ -3 \end{bmatrix},$$

and the parametric equations as:

$$\begin{cases} x_1 &= 1 + t \\ x_2 &= 2 - 4t \\ x_3 &= 3 - 3t \end{cases}.$$

■

Example 2.4.7. Find a vector equation and parametric equations of the plane in $\mathbb{R}^4$ that passes through $(-17, 6, 29, 0)$, $(-13, 3, 25, -2)$, and $(-15, 6, 25, -1)$.

We can take $\boldsymbol{x}_0$ to be any of the three points. We will take $\boldsymbol{x}_0 = (-17, 6, 29, 0)$. To get the spanners $\boldsymbol{u}$ and $\boldsymbol{v}$, we need to take the differences of the three vectors provided, that is,

$$\boldsymbol{u} = (-13, 3, 25, -2) - (-17, 6, 29, 0) = (4, -3, -4, -2), \quad \text{and} \quad \boldsymbol{v} = (-15, 6, 25, -1) - (-17, 6, 29, 0) = (2, 0, -4, -1).$$

Therefore, the vector equation is given as:

$$\boldsymbol{x} = \boldsymbol{x_0} + a\boldsymbol{u} + b\boldsymbol{v},$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -17 \\ 6 \\ 29 \\ 0 \end{bmatrix} + a \begin{bmatrix} 4 \\ -3 \\ -4 \\ -2 \end{bmatrix} + b \begin{bmatrix} 2 \\ 0 \\ -4 \\ -1 \end{bmatrix},$$

and the parametric equations as

$$\begin{cases} x_1 &= -17 + 4a + 2b \\ x_2 &= 6 - 3a \\ x_3 &= 29 - 4a - 4b \\ x_4 &= -2a - b \end{cases}.$$

■

Note that each flat corresponds to the solution set of a linear system: the parametric equations for the *Bottom-Up* approach or the $m \times n$ linear system for the *Top-Down* approach. In the *Top-Down* approach, to compute the intersection of two flats in F^n, one need only take the union of all the linear equations from their associated linear systems and solve the combined linear system. This really is the drive between *Top-Down*: intersections! For computing the intersection with the *Bottom-Up* approach, it is important to remember that we use different symbols for the free variables in the different parametric equations. The intersection of flats is the set of all points that are in both subsets. The parameters that give these coincident points do not have to be the same. If the same symbols for the parameters are used, it makes the false impression that the parameters must be equal, which is not true in general.

Example 2.4.8. Find the intersection for the two planes in $\mathbb{R}^4$ from Examples 2.4.4 and 2.4.7.

To begin, we can equate the values for the dependent variables x_1, x_2, x_3, x_4, as follows:

$$\begin{cases} 26 + s &= -17 + 4a + 2b \\ 3 - 3s &= 6 - 3a \\ -13 - 2s + t &= 29 - 4a - 4b \\ -18 - s &= -2a - b \end{cases} \sim \begin{cases} 4a + 2b - s &= 43 \\ -3a + 3s &= -3 \\ -4a - 4b + 2s - t &= -42 \\ -2a - b + s &= -18 \end{cases} \sim \begin{cases} a &= 8 \\ b &= 9 \\ s &= 7 \\ t &= -12 \end{cases}$$

Solving this linear system gives the values $a = 8$ and $b = 9$ ($s = 7$ and $t = -12$), which implies that there is a unique point of intersection between the two planes, $\boxed{(33, -18, -39, -25)}$. As bizarre as it may be to visualize, although two distinct planes cannot intersect at a unique point in 3-space, they can in 4-space. ■

Example 2.4.9. In the following example, we will compute the intersection between two affine sets, but we will do it twice! The first time we will compute the intersection when the flats are represented *Top-Down*, and the second attempt will be when the flats are represented *Bottom-Up*. We present them side-by-side for

the reader to compare the differences in the representation.

Find the intersection of the two affine sets in $\mathbb{R}^3$ given as:

Top-Down

$$\begin{cases} x_1 + 3x_2 - x_3 = 16 \\ x_1 + 4x_2 + x_3 = 24 \end{cases}$$

and

$$\begin{cases} 3x_1 + 7x_2 - 6x_3 = 34 \\ -3x_1 - 8x_2 + 4x_3 = -42. \end{cases}$$

To begin, we combine the two linear systems into one. Note that any vector on the first flat is a solution to the first two equations, and any vector on the second flat is a solution to the second two equations. Hence, a vector on their intersection will be a solution to all four linear equations.

$$\begin{cases} x_1 + 3x_2 - x_3 = 16 \\ x_1 + 4x_2 + x_3 = 24 \\ 3x_1 + 7x_2 - 6x_3 = 34 \\ -3x_1 - 8x_2 + 4x_3 = -42. \end{cases}$$

We proceed to solve the linear system:

$$\left[\begin{array}{ccc|c} 1 & 3 & -1 & 16 \\ 1 & 4 & 1 & 24 \\ 3 & 7 & -6 & 34 \\ -3 & -8 & 4 & -32 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Therefore, the intersection of the two affine sets is the point $\boldsymbol{x} = \boxed{(6, 4, 2)}$.

Bottom-Up

$$\begin{cases} x_1 = -1 + 7s \\ x_2 = 6 + -2s \\ x_3 = 1 + s \end{cases} \text{ and } \begin{cases} x_1 = -14 + 20t \\ x_2 = 10 - 6t \\ x_3 = -1 + 3t \end{cases}$$

To begin, we equate the variables for each of the parametric equations. Then remove the coordinate vector and solve the parameters as a linear system.

$$\begin{cases} 7s - 1 = x_1 = 20t - 14 \\ -2s + 6 = x_2 = -6t + 10 \\ s + 1 = x_3 = 3t - 1 \end{cases} \sim \begin{cases} 7s - 20t = -13 \\ -2s + 6t = 4 \\ s - 3t = -2 \end{cases}$$

We proceed to solve the linear system:

$$\left[\begin{array}{cc|c} 7 & -20 & -13 \\ -2 & 6 & 4 \\ 1 & -3 & -2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right]$$

Hence, the intersection occurs when the parameters are $s = 1$ and $t = 1$. Therefore, the intersection of the two affine sets is the point $\boldsymbol{x} = (-1 + 7(1), 6 - 2(1), 1 + (1)) = (-14 + 20(1), 10 - 6(1), -1 + 3(1)) = \boxed{(6, 4, 2)}$. ∎

Exercises

(Go to Solutions)

For Exercises 1-6, describe what vector space the given affine set resembles. Assume all vectors and equations are linearly independent.

♠ 1. A linear system with 4 equations and 6 variables.

♠ 2. A linear system with 7 equations and 8 variables.

♠ 3. A linear system with 4 equations and 12 variables.

♠ 4. The span of three vectors.

♠ 5. The translations of the span of five vectors.

♠ 6. The span of no vectors, that is, Span{}.

For Exercises 7-14, find the vector form and parametric equations for the given affine set. Answers may vary.

7. The line passing through $(5, 4, 6)$ and $(1, -2, -3)$ in $\mathbb{R}^3$.

8. The plane containing $(-3, 6, 1)$ and parallel to $(4, -2, 3)$ and $(1, 6, 5)$ in $\mathbb{R}^3$.

9. The plane containing $(1, 2, 3)$, $(-1, 4, 1)$, $(4, -2, 8)$, $(0, 1, 6)$ in $\mathbb{R}^3$.

♠ 10. The plane containing $(1, 2, 3, 4)$, $(0, 1, 0, 1)$, and $(-1, -2, 1, 4)$ in $\mathbb{R}^4$.

♠ 11. The hyperplane containing $(0, 1, 4, -1)$, $(11, 0, 0, 2)$, $(-2, -3, 0, 1)$, and $(9, 9, 0, -1)$ in $\mathbb{R}^4$.

♠ 12. The line containing $(i, 1 + 2i)$ and parallel to the vector $(2 - i, 5)$ in $\mathbb{C}^2$.

♠ 13. The line passing through $(1, 2, 3)$ and $(0, 1, 3)$ in $\mathbb{Z}_5^3$.

♠ 14. The plane containing $(2, 3, 1)$ and parallel to $(1, 2, 3)$ and $(0, 0, 5)$ in $\mathbb{Z}_7^3$.

For Exercises 15-19, find the intersection between the two flats provided. Answers may vary.

15. $\begin{cases} x_1 + x_2 + x_3 - 5x_4 = 3 \\ 2x_1 + 3x_2 + 2x_4 = -4 \end{cases}$ and $\begin{cases} 5x_1 + 2x_2 - x_3 - x_4 = -3 \\ -3x_1 + 2x_3 - 4x_4 = 5 \end{cases}$

♠ 16. $\begin{cases} -x_1 - 6x_3 - x_4 = -17 \\ -3x_1 + 2x_2 - 10x_3 - 3x_4 = -33 \end{cases}$ and $\begin{cases} x_1 - 2x_2 - 2x_3 + x_4 = -1 \\ x_2 + 3x_3 - 3x_4 = 5 \end{cases}$

17. $\boldsymbol{x} = \begin{bmatrix} -2 \\ 2 \\ -4 \end{bmatrix} + a \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 4 \\ -2 \\ 0 \end{bmatrix} + c \begin{bmatrix} -3 \\ 1 \\ 4 \end{bmatrix}$

♠ 18. $\boldsymbol{x} = \begin{bmatrix} 4 \\ 2 \\ 6 \\ -2 \end{bmatrix} + a \begin{bmatrix} -1 \\ 0 \\ -1 \\ 2 \end{bmatrix}$ and $\boldsymbol{x} = \begin{bmatrix} 4 \\ -2 \\ 4 \\ 15 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 3 \\ -5 \end{bmatrix} + c \begin{bmatrix} 2 \\ -2 \\ 2 \\ 3 \end{bmatrix}$.

19. The affine sets from Exercises 7 and 8.

"For the wise man looks into space and he knows there is no limited dimensions." – Lao Tzu

2.5 Subspaces

Lines, planes, and hyperplanes through the origin are just special examples of subspaces of F^n.

Definition 2.5.1. Let V be a vector space, such as F^n. A **subspace** of V is any set $W \subseteq V$ such that:

(i) $\mathbf{0} \in W$,

(ii) For each $\boldsymbol{u}, \boldsymbol{v} \in W$, the sum $\boldsymbol{u} + \boldsymbol{v} \in W$,

(iii) For each $\boldsymbol{v} \in W$ and scalar $c \in F$, the vector $c\boldsymbol{v} \in W$.

In words, a subspace is nonempty subset of V *closed* under addition and scalar multiplication. It should be mentioned that closure under addition and scalar multiplication is equivalent to closure under linear combinations.

When considering vector spaces before, we saw that a set of things is called a vector space when we can appropriately add and scale all the objects, which are then called vectors. Necessary to this definition is that the sum of two vectors and a scaled vector still be vectors! A subspace is a vector space inside of a vector space, that is, each subspace of V is also a vector space in its own right. The sum of two vectors or a scaled vectors originating from W will remain in W still. In regard to the axioms of a vector space, the subspace inherits the axioms from the ambient vector space. For example, since ALL vectors commute in the larger vector space, all vectors will commute in the smaller subset.

The vector space V in consideration will nearly always be F^n for some field F. It should be mentioned that F^n is a subspace of F^n because it satisfies all three conditions. Also, the **zero space** $F^0 = \{\mathbf{0}\}$ is also a subspace.

Example 2.5.2. If $\boldsymbol{v}_1, \boldsymbol{v}_2 \in F^n$ and $W = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, then we claim that W is a subspace of F^n. To prove this claim, we must show that W satisfies the three conditions in the definition. Let $c, s_1, s_2, t_1, t_2 \in F$. Then

(i) $\mathbf{0} = 0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 \in W$,

(ii) $(s_1\boldsymbol{v}_1 + s_2\boldsymbol{v}_2) + (t_1\boldsymbol{v}_1 + t_2\boldsymbol{v}_2) = (s_1 + t_1)\boldsymbol{v}_1 + (s_2 + t_2)\boldsymbol{v}_2 \in W$,

(iii) $c(s_1\boldsymbol{v}_1 + s_2\boldsymbol{v}_2) = (cs_1)\boldsymbol{v}_1 + (cs_2)\boldsymbol{v}_2 \in W$.

Therefore, W is a subspace of F^n. By mathematical induction, this same argument shows that a span for any number of vectors is also a subspace of F^n. ∎

Because of this, we often say that a span of a set of vector $\mathcal{S}$ is the **subspace spanned by** $\mathcal{S}$. Likewise, if H is a subspace of F^n and $W = \text{Span}\{\mathcal{S}\}$ for some $\mathcal{S} \subseteq F^n$, then we say that $\mathcal{S}$ is a **spanning set** for W.

By our consideration of lines, planes, hyperplanes, and affine sets (flats) which pass through the origin are subspaces because it is just a span of vectors. In fact, every coset is simply the translation of a subspace. On the other hand, a translation of a subspace will likely not be a subspace since it no longer contains the zero vector.

Example 2.5.3. In $\mathbb{R}^2$, all lines through the origin are 1-dimensional subspaces. On the other hand, other lines are not subspaces since they do not contain the origin which is the zero vector of $\mathbb{R}^2$.

Let W be the set of all points $(x, y) \in \mathbb{R}^2$ for which $x \geq 0$ and $y \geq 0$, that is, W is the first quadrant of the plane. This is not a subspace. On the one hand, the set contains $\mathbf{0}$ and is closed under addition. For example, if $x_1, x_2, y_1, y_2 \geq 0$, then $x_1 + x_2, y_1 + y_2 \geq 0$ and $(x_1, y_2) + (x_2, y_2) \in W$. On the other hand, if $x, y \geq 0$, then $-x, -y \leq 0$ and $-(x, y) = (-x, -y) \notin W$. Hence, W is not a subspace of $\mathbb{R}^2$. ∎

Let $A = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{bmatrix}$ be an $m \times n$ matrix. Then recall that the **column space** of A, denoted $\operatorname{Col} A$, is $\operatorname{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}$. The column space of A is the set of all $\boldsymbol{b} \in F^m$ such that $A\boldsymbol{x} = \boldsymbol{b}$ is consistent. By Example 2.5.2, $\operatorname{Col} A$ is a subspace of F^m.

Example 2.5.4. Let $A = \begin{bmatrix} 1 & 8 & 7 \\ 7 & 6 & 9 \\ 8 & 7 & 6 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 3 \\ 3 \\ 7 \end{bmatrix}$ over $\mathbb{Z}_{11}$. Determine whether $\boldsymbol{b} \in \operatorname{Col} A$.

We need to determine whether $\boldsymbol{b}$ is a linear combination of the column vectors of A. This is the same as solving the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$. To do this, we use row reduction:

$$\left[\begin{array}{ccc|c} 1 & 8 & 7 & 3 \\ 7 & 6 & 9 & 3 \\ 8 & 7 & 6 & 7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 8 & 7 & 3 \\ 0 & 5 & 4 & 4 \\ 0 & 9 & 5 & 5 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 8 & 7 & 3 \\ 0 & 5 & 4 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right] \pmod{11}.$$

Thus, the system is consistent, which implies that $\boldsymbol{b}$ is in $\operatorname{Col} A$. ∎

Example 2.5.5.

(a) Let $V = \mathbb{R}^\infty$ be the set of all real-valued sequences. This is a vector space since we can add two sequences:

$$\{x_1, x_2, x_3, \dots\} + \{y_1, y_2, y_3, \dots\} = \{x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots\} \in \mathbb{R}^\infty,$$

and scale a sequence:

$$c\{x_1, x_2, x_3, \dots\} = \{cx_1, cx_2, cx_3, \dots\}.$$

Notice the zero vector of this vector space is the constant **zero sequence**:

$$\mathbf{0} = \{0, 0, 0, \dots\}.$$

Let W be the subset of $\mathbb{R}^\infty$ of convergent sequences. Since the zero sequence converges (to zero), it is contained in W. By limit properties, the sum of convergent sequences is convergent and converges to the sum of limits. Likewise, a multiple of a convergent sequence converges to a multiple of the limit. Therefore, W is a subspace of $\mathbb{R}^\infty$. By similar reasoning, the set of all sequence which converge to zero is a subspace of $\mathbb{R}^\infty$.

(b) Let $\mathbb{R}^X = \{f : X \to \mathbb{R}\}$, that is, the set of all real-valued functions where $X \subseteq \mathbb{R}$, e.g. $\mathbb{R}^\mathbb{R}$ is the set of real-valued function defined on the entire x-axis. Then $\mathbb{R}^X$ is also a vector space. Note that if $f, g \in \mathbb{R}^X$, then the sum of "vectors" is the function defined by the rule:

$$(f + g)(x) = f(x) + g(x),$$

and scalar multiplication is given by the rule:

$$(cf)(x) = c[f(x)].$$

For this vector space, the function $f(x) = 0$ for all $x \in X$ is the zero vector, called the **zero function**.

Let $\mathcal{P}$ denote the set of all polynomials with real coefficients. Viewing $\mathcal{P}$ as a subset of $\mathbb{R}^X$, as the domain of any polynomial can be restricted to X, then $\mathcal{P}$ is a subspace of $\mathbb{R}^\mathbb{R}$. To see this, note that the **zero polynomial** is the same as the zero function and is contained in $\mathcal{P}$. Second, the sum of two

polynomials is again a polynomial. Third, a polynomial times by a real number is again a real-valued polynomial. Therefore, $\mathcal{P}$ is a subspace of $\mathbb{R}^{\mathbb{R}}$.

Let $\mathcal{P}_n$ be the set of polynomials with degree at most n. By the same reasoning as before, $\mathcal{P}_n$ is a subspace of $\mathcal{P}$ and hence a subspace of $\mathbb{R}^X$. In fact, $\mathcal{P}_n = \text{Span}(1, x, x^2, \ldots, x^n)$. Furthermore,

$$\mathcal{P}_0 \leq \mathcal{P}_1 \leq \mathcal{P}_2 \leq \mathcal{P}_3 \leq \ldots \leq \mathcal{P} \leq \mathbb{R}^X.$$

(c) Let $\mathcal{C}(X)$ be the set of all real-valued continuous functions on the domain X. This is a subset of $\mathbb{R}^X$. By facts from Calculus, the zero function is continuous, the sum of continuous functions is continuous, and the multiple of a continuous function is continuous. Thus, $\mathcal{C}(X) \leq \mathbb{R}^X$.

Likewise, we can define the set $\mathcal{C}^1(X)$ to be the set of all real-valued continuously differentiable (f' is continuous) functions on the domain X. Likewise, Calculus tells us that constant functions are differentiable, sums of differentiable functions are differentiable, and multiples of differentiable functions are differentiable. Thus, $\mathcal{C}^1(X) \leq \mathbb{R}^X$. Of course, since differentiable functions are necessarily continuous, we have $\mathcal{C}^1(X) \leq \mathcal{C}(X) \leq \mathbb{R}^X$.

Likewise, we can define $\mathcal{C}^n(X)$ to be the set of functions f for which $f^{(n)}$ is continuous on X. By similar reasoning, each of these sets are subspaces of $\mathbb{R}^X$. Let $\mathcal{C}^\infty(X)$ be the set of all functions in $\mathbb{R}^X$ for which all higher derivatives exists (and necessarily are continuous). This is likewise a subspace and called the space of **smooth functions**. In fact, we have the following descending sequences of subspaces:

$$\begin{aligned} \mathbb{R}^X &\geq \mathcal{C}(X) \geq \mathcal{C}^1(X) \geq \mathcal{C}^2(X) \geq \ldots \geq \mathcal{C}^n(X) \geq \ldots \geq \mathcal{C}^\infty(X) \\ &\geq \mathcal{P} \geq \ldots \geq \mathcal{P}_n \geq \ldots \geq \mathcal{P}_2 \geq \mathcal{P}_1 \geq \mathcal{P}_0. \end{aligned}$$

All these vector spaces allude to why Linear Algebra is extremely useful for Advanced Calculus (often called Real Analysis). Linear algebra is everywhere in calculus!

With the exception of those concepts directly related to calculus, such as continuity and derivatives, the above examples of function spaces can be adapted to any field F. ∎

Exercises

(Go to Solutions)

For Exercises 1-6, explain why the subset of $\mathbb{R}^2$ does NOT form a subspace, that is, explain which of the three axioms are satisfied and which ones fail. Provide a counterexample to justify when they fail. Answers may vary.

♠ 1. W is the unit circle, that is, $W = \{(x, y) \mid x^2 + y^2 = 1\}$.

♠ 2. W is the unit disc, that is, $W = \{(x, y) \mid x^2 + y^2 \leq 1\}$.

♠ 3. W is the standard parabola, that is, $W = \{(x, y) \mid y = x^2\}$.

♠ 4. W is the union of the x- and y-axis, that is, $W = \{(x, y) \mid x = 0 \text{ or } y = 0\}$.

♠ 5. W is the upper half-plane, that is, $W = \{(x, y) \mid y \geq 0\}$.

6. W is the punctured plane, that is, $W = \{(x, y) \mid x^2 + y^2 \geq 1\}$.

For Exercises 7-10, decide if the subset of the function space $\mathbb{R}^{\mathbb{R}}$ forms a subspace. Justify your answer in a similar style to Exercises 1-6. Answers may vary.

♠ 7. The set of functions whose y-intercept is 1, that is, $f(0) = 1$.

♠ 8. The set of functions whose y-intercept is 0, that is, $f(0) = 0$.

♠ 9. The set of functions whose end behavior on the right is ∞, that is, $\lim_{x \to \infty} f(x) = \infty$.

♠ 10. The set of odd functions, that is, $f(-x) = -f(x)$.

"After every storm the sun will smile; for every problem there is a solution, and the soul's indefeasible duty is to be of good cheer." – William R. Alger

2.6 Solution Sets of Linear Systems

Recall that a linear system is said to be *homogeneous* if it can be written in the form $A\boldsymbol{x} = \mathbf{0}$, where A is a $m \times n$ matrix, $\boldsymbol{x} \in F^n$, and $\mathbf{0}$ is the zero vector in F^m. A homogeneous system $A\boldsymbol{x} = \mathbf{0}$ always has a solution, namely $\boldsymbol{x} = \mathbf{0}$. This can be thought of as the **trivial solution**. Any other solution to a homogeneous system is a **nontrivial solution**.

Theorem 2.6.1. *The homogeneous system $A\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution if and only if the equation has at least one free variable.*

Example 2.6.2. Determine if the following homogeneous system has a nontrivial solution.

$$\begin{cases} 3x_1 + 5x_2 + 3x_3 \equiv 0 \\ 4x_1 + 5x_2 + 4x_3 \equiv 0 \pmod 7. \\ 6x_1 + x_2 + 6x_3 \equiv 0 \end{cases}$$

Let A be the coefficient matrix. Then we must row reduce the augmented matrix $[A \mid \mathbf{0}]$, as seen below.

$$[A \mid \mathbf{0}] \equiv \left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 4 & 5 & 4 & 0 \\ 6 & 1 & 6 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 5 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, x_3 is a free variable. This implies that $A\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution! To determine such a solution, we continue to the RREF;

$$\left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 0 & 3 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, the system of equations becomes:

$$\begin{cases} x_1 + x_3 = 0 \\ x_2 = 0 \pmod 7 \\ 0 = 0 \end{cases}$$

and the nontrivial solutions are of the form:

$$\boldsymbol{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} 6x_3 \\ 0 \\ x_3 \end{bmatrix} \equiv x_3 \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix} \pmod 7.$$

Thus, if $\boldsymbol{v} \equiv \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$, then the solution set of $A\boldsymbol{x} = \mathbf{0}$ is

$$\operatorname{Span}\{\boldsymbol{v}\} = \{t\boldsymbol{v} \mid t \in \mathbb{Z}_7\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 6 \end{bmatrix} \right\}. \qquad \blacksquare$$

Definition 2.6.3. Let A be an $m \times n$ matrix. Then the **null space** (or **kernel**) of A, denoted $\operatorname{Nul} A$, is the set of all solutions of the homogeneous system $A\boldsymbol{x} = \mathbf{0}$. The **nullity** of A, denoted $\operatorname{nullity}(A)$, is the number of non-pivot columns in A. This value counts the number of free variables in the homogeneous system and essentially measures the "size" of null space.

The null space A is a subset of F^n. Furthermore, $\operatorname{Nul} A$ is a subspace. This agrees with the previous example where the solution set was spans of vectors. As such, we can geometrically visualize the solution set to a homogeneous system as a plane through the origin.

Theorem 2.6.4. *The null space of an $m \times n$ matrix A is a subspace of F^n.*

Proof. Since $A\mathbf{0} = \mathbf{0}$, we have that $\mathbf{0} \in \operatorname{Nul} A$. Next, suppose that $\boldsymbol{x}, \boldsymbol{y} \in \operatorname{Nul} A$. This means that $A\boldsymbol{x} = \mathbf{0}$ and $A\boldsymbol{y} = \mathbf{0}$. Thus, $A(\boldsymbol{x} + \boldsymbol{y}) = A\boldsymbol{x} + A\boldsymbol{y} = \mathbf{0} + \mathbf{0} = \mathbf{0}$. Hence, $\boldsymbol{x} + \boldsymbol{y} \in \operatorname{Nul} A$. Finally, if $c \in F$, then $A(c\boldsymbol{x}) = c(A\boldsymbol{x}) = c\mathbf{0} = \mathbf{0}$. Hence, $c\boldsymbol{x} \in \operatorname{Nul} A$. Therefore, $\operatorname{Nul} A$ is a subspace of F^n. □

Example 2.6.5. For the matrix $A = \begin{bmatrix} 1 & 3 & -3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & -2 & 13 & 1 \end{bmatrix}$, find a spanning set for its null space.

Note that when solving a homogeneous linear system, the augmented column is just the zero vector and no row operation will ever transform a column of zeros. As such, we often omit the zero column when solving homogeneous systems. In fact, solving homogeneous systems is essentially the same techniques as determining whether a set of vectors is linearly independent (compare Example 2.3.2).

$$\begin{bmatrix} 1 & 3 & -3 & 4 \\ 2 & 1 & 4 & 3 \\ 3 & -2 & 13 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 4 \\ 0 & -5 & 10 & -5 \\ 0 & -11 & 22 & -11 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & 1 & -2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & -3 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 3 & 1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The associated reduced linear system would then be

$$\begin{cases} x_1 + 3x_3 + x_4 = 0 \\ x_2 - 2x_3 + x_4 = 0 \\ 0 = 0 \end{cases} \sim \begin{cases} x_1 = -3x_3 - x_4 \\ x_2 = 2x_3 - x_4 \end{cases}$$

Thus, the general solution to the homogeneous system $A\boldsymbol{x} = \mathbf{0}$ is

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_3 - x_4 \\ 2x_3 - x_4 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3x_3 \\ 2x_3 \\ x_3 \\ 0 \end{bmatrix} + \begin{bmatrix} -x_4 \\ -x_4 \\ 0 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -3 \\ 2 \\ 1 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 1 \end{bmatrix}.$$

Let $\boldsymbol{u} = (-3, 2, 1, 0)$ and $\boldsymbol{v} = (-1, -1, 0, 1)$. Hence, $\text{Nul}(A) = \text{Span}\{\boldsymbol{u}, \boldsymbol{v}\}$. ∎

Like a homogeneous system, a non-homogeneous system of linear equations with multiple solutions can have their solutions expressed in parametric form.

Example 2.6.6. Describe all solutions of $A\boldsymbol{x} = \boldsymbol{b}$ over $\mathbb{Z}_7$, where

$$A = \begin{bmatrix} 3 & 5 & 3 \\ 4 & 5 & 4 \\ 6 & 1 & 6 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} 0 \\ 6 \\ 3 \end{bmatrix}.$$

Here, A is the same coefficient matrix as in the first example. Thus, using the same row operations, we can see that

$$\left[\begin{array}{ccc|c} 3 & 5 & 3 & 0 \\ 4 & 5 & 4 & 6 \\ 6 & 1 & 6 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 1 & 6 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, the related system is

$$\begin{cases} x_1 + x_3 = 6 \\ x_2 = 2 \\ 0 = 0. \end{cases}$$

and the solutions are all of the form:

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 + 6x_3 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}.$$

Thus, if $\boldsymbol{v} = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$ and $\boldsymbol{x}_0 = \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}$, then all solutions of $A\boldsymbol{x} = \boldsymbol{b}$ are points on the line determined by $\boldsymbol{x} = \boldsymbol{x}_0 + t\boldsymbol{v}$, $t \in \mathbb{Z}_7$. In fact, $t\boldsymbol{v}$ is the general form of the solutions for $A\boldsymbol{x} = \mathbf{0}$. Thus, any solution to $A\boldsymbol{x} = \boldsymbol{b}$ is of the form $\boldsymbol{x}_0$ plus a solution to $A\boldsymbol{v} = \mathbf{0}$. Using the $\text{Span}\{\boldsymbol{v}\}$ from Example 2.6.2, we see the seven solutions are

$$\text{Aff}\left\{\begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}\right\} = \left\{\begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 6 \end{bmatrix}\right\}.$$ ∎

Theorem 2.6.7. *Suppose the equation $A\boldsymbol{x} = \boldsymbol{b}$ is consistent for some given $\boldsymbol{b}$, and let $\boldsymbol{x}_0$ be a (particular) solution. Then the solution set of $A\boldsymbol{x} = \boldsymbol{b}$ is the set of all vectors of the form $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_n$, where $\boldsymbol{x}_n$ is any solution of the homogeneous equation $A\boldsymbol{x} = \boldsymbol{0}$ and can be expressed in parametric form.*

From a geometric perspective, a solution set to a homogeneous system is the null space of its coefficient matrix A. As such, it is a subspace spanned by n independent vectors, where $n = \text{nullity}(A)$. Theorem 2.6.7 tells us that a solution set to a non-homogeneous system generically is an affine span of n independent vectors, in other words, it is an n-flat.

Exercises

(Go to Solutions)

For Exercises 1-1, determine whether the homogeneous system has a nontrivial solution. If so, list all the free variables, e.g., x_2, x_4, x_5.

1. $\begin{cases} 2x_1 + 3x_2 + 2x_3 = 0 \\ 2x_1 + 5x_2 + 7x_3 = 0 \\ 4x_1 + 6x_2 + 4x_3 = 0 \\ -2x_1 + x_2 + 8x_3 = 0 \end{cases}$

For Exercises 2-8, find a spanning set for the null space for each of the following matrices. Answers may vary.

2. $\begin{bmatrix} 9 & 12 & -9 & 3 \\ 10 & 11 & -4 & 2 \\ 4 & 7 & 2 & 0 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 2 & -4 & 0 \\ -2 & 4 & 1 \\ 1 & -2 & 3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & -2 & -2 \\ 1 & -1 & -4 & -3 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 & 2 \\ -2 & -4 \\ 3 & 6 \\ 4 & 8 \end{bmatrix}$

♠ 6. $\begin{bmatrix} -5 & 10 & 7 & 4 & -30 \\ -1 & 2 & 1 & 0 & -8 \\ -1 & 2 & 2 & 2 & -3 \\ 4 & -8 & -3 & 0 & 29 \\ 1 & -2 & 2 & -1 & -5 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 2 & -1 & -4 & 10 & 10 \\ 1 & 0 & -1 & 3 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 0 & 2 & -1 & -2 \\ 2 & 1 & -1 & -2 & 1 \\ 4 & 2 & -2 & -4 & 2 \end{bmatrix}$

For Exercises 9-13, describe all solutions to the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$. The matrix A will be listed first, followed by the vector $\boldsymbol{b}$. Answers may vary.

9. $\begin{bmatrix} 2 & -6 & 8 & 0 \\ 1 & -3 & 4 & 1 \\ 0 & 0 & 0 & 1 \\ -1 & 3 & -4 & 1 \end{bmatrix}, \begin{bmatrix} 10 \\ -3 \\ -8 \\ -13 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 9 & -4 & -6 \\ -1 & 6 & -16 \\ -1 & -2 & 8 \end{bmatrix}, \begin{bmatrix} 1 \\ 11 \\ -5 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 2 & -2 & 2 & 0 \\ 3 & 4 & -5 & 0 \\ 0 & 0 & 6 & 0 \\ -2 & -1 & 4 & 0 \end{bmatrix}, \begin{bmatrix} 4 \\ -4 \\ 18 \\ 8 \end{bmatrix}$

♠ 12. $\begin{bmatrix} 3-4i & 3+i & 6+2i \\ 2 & i & 2i \end{bmatrix}, \begin{bmatrix} 12+14i \\ -2+5i \end{bmatrix}$

♠ 13. $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \pmod 2$

We have seen that solution sets to homogeneous systems, aka null spaces, are subspaces, that is, they always contain the zero vector, are always closed under addition, and are always closed under scalar multiplication. In particular, if one has a list of solutions to the homogeneous system, then any linear combination of these vectors is likewise a solution to the system. What about solution sets of non-homogeneous systems?

After all, the solution set to a linear system is a flat and subspaces are just flats through the origin. Sadly, such closure principles do not hold for general flats. Firstly, if the flat does not pass through the origin, then it cannot be a subspace since it does not contain the zero vector. Similar problems arise with closure under addition and scalars.

♠ 14. Construct an example of a non-homogeneous system such that the sum of two solutions is not a solution.

♠ 15. Construct an example of a non-homogeneous system such that the scalar multiple of a solution is not a solution.

♠ 16. Although solution sets of non-homogeneous systems are not closed under vector addition or scalar multiples, they are closed under lines, that is, if $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are two distinct solutions to the non-homogeneous system then any vector on the line

$$\boldsymbol{x} = (1-t)\boldsymbol{x}_0 + t\boldsymbol{x}_1, \quad t \in F$$

is likewise a solution to the system. Prove that solution sets are closed under lines.

17. We have already alluded to the fact that flat and affine spans are actually the same thing, despite being defined differently. In this direction, prove that a solution set is closed under all affine combinations.

"It is through gratitude for the present moment that the spiritual dimension of life opens up."
– Eckhart Tolle

2.7 Bases

Definition 2.7.1. A **basis** $\mathcal{B}$ of a vector space V is a linearly independent, spanning set of V. The size of the basis, $|\mathcal{B}|$, is called the **dimension** of V, denoted $\dim V$.

Every line through the origin is a one-dimensional subspace and every plane through the origin is two-dimensional. In general, a flat through the origin spanned by p independent vectors (aka, a subspace) has dimension p. Consider for a moment the zero space $\{\mathbf{0}\}$. This subspace has exactly two subsets $\{\mathbf{0}\}$ and $\emptyset$. Vacuously, $\emptyset$ is linearly independent[†]. The set $\{\mathbf{0}\}$ is linearly dependent because it contains $\mathbf{0}$. Thus, $\emptyset$ is a basis for $\{\mathbf{0}\}$. Therefore, $\dim\{\mathbf{0}\} = 0$.

Example 2.7.2. For F^n, the set $\mathcal{E} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$, where $\boldsymbol{e}_i$ is the vector with a 1 in the ith component and 0's everywhere else, is always a basis. This is known as the **standard basis** of F^n.

On the other hand, the set $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is a non-standard basis for F^3. ∎

We mention some important properties of bases and dimension.

Theorem 2.7.3. *Let $\mathcal{B}$ be a subset of the vector space V. Then the following are equivalent:*

(i) $\mathcal{B}$ is a basis;

(ii) $\mathcal{B}$ is a maximal linearly independent set, that is, $\mathcal{B}$ is linearly independent but $\mathcal{B} \cup \{\boldsymbol{u}\}$ is linearly dependent for any $\boldsymbol{u} \in V$;

(iii) $\mathcal{B}$ is a minimal spanning set of V, that is, $V = \operatorname{Span}(\mathcal{B})$ but $V \neq \operatorname{Span}(\mathcal{B} \setminus \{\boldsymbol{v}\})$ for any $\boldsymbol{v} \in \mathcal{B}$.

Theorem 2.7.4 (The Expansion Theorem)**.** *Let V be a vector space and let $S \subseteq V$ be a linearly independent subset. Then there exists a basis B of V such that $S \subseteq B$.*

In other words, every linearly independent set can be expanded into a basis by including potentially new vectors from V not already contained in S.

Theorem 2.7.5 (The Pruning Theorem)**.** *Let V be a vector space and let $S \subseteq V$ be a spanning set of V. Then there exists a basis B of V such that $B \subseteq S$.*

In other words, every spanning set can be pruned down into a basis by removing unneeded vectors from S.

Theorem 2.7.6. *The dimension of a vector space V is well-defined, that is, if $\mathcal{B}$ and $\mathcal{C}$ are two bases for V, then $|\mathcal{B}| = |\mathcal{C}|$.*

Theorem 2.7.7. *Let V be an n-dimensional vector space and let $S \subseteq V$.*

(i) If $|S| > n$, then S is linearly dependent. But if S spans V, then there is a proper subset of S which is a basis.

(ii) If $|S| < n$, then S does not span V. But if S is linearly independent, then there is a basis of V that contains S as a proper subset.

Given any $m \times n$ matrix A, there are two fundamental subspaces[††] associated to A: the *column space* $\operatorname{Col}(A)$ which is the span of the column vectors of A (hence, a subspace of F^m) and the *null space* $\operatorname{Nul}(A)$ which is the solution set to the homogeneous system $A\boldsymbol{x} = \mathbf{0}$ (hence, a subspace of F^n). The dimension of $\operatorname{Col}(A)$ is called the **rank**[‡] of A, and the dimension of $\operatorname{Nul}(A)$ is called the **nullity** of A.

Theorem 2.7.8. *The pivot columns of A form a basis for the column space of A.*

The location of the pivots (or to say, the absence of pivots) in the echelon form of A tells us which column vectors can be expressed as a linear combination of the previous column vectors

Example 2.7.9. Let $A = \begin{bmatrix} 1 & 3 & 3 & 2 & 4 \\ 2 & 7 & 6 & 3 & 9 \\ 1 & 2 & 3 & 3 & 3 \end{bmatrix}$. Construct a basis for $\mathrm{Col}(A)$.

We begin by computing an echelon form of A:

$$\begin{bmatrix} \boxed{1} & 3 & 3 & 2 & 4 \\ 2 & 7 & 6 & 3 & 9 \\ 1 & 2 & 3 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 3 & 3 & 2 & 4 \\ 0 & \boxed{1} & 0 & -1 & 1 \\ 0 & -1 & 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 3 & 3 & 2 & 4 \\ 0 & \boxed{1} & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

We see now that this matrix is in echelon form, for which we can now easily identify the pivot columns, namely the first and second. This indicates that the third and fourth column vectors are both linear combinations of the first two column vectors. Therefore, a basis of $\mathrm{Col}\, A$ is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ 2 \end{bmatrix} \right\}$ and $\mathrm{rank}(A) = 2$. ∎

Example 2.7.10. Let $A = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix}$. Find a basis for the column space of A.

Since $A \sim \begin{bmatrix} \boxed{1} & 0 & -3 & 5 & 0 \\ 0 & \boxed{1} & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, we see that the third and fourth column vectors are linear combinations of the first and second column vectors. Thus, the set $\left\{ \begin{bmatrix} 1 \\ -2 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -9 \\ 2 \\ 1 \\ -8 \end{bmatrix} \right\}$ is a basis for $\mathrm{Col}\, A$, and the $\mathrm{rank}(A) = \dim(\mathrm{Col}(A)) = 3$. ∎

Theorem 2.7.11. *If two matrices A and B are row equivalent, then* $\mathrm{Nul}\, A = \mathrm{Nul}\, B$.

Because of the previous theorem, one can use the row-reduced echelon form of A to find a basis for $\mathrm{Nul}(A)$. This will be the the same technique we applied in Example 2.6.5 to find a spanning set for $\mathrm{Nul}(A)$, which was already a basis. The basis vectors will be derived from the non-pivot columns of A's row-reduced

echelon form, that is, those columns which correspond to free variables of the linear system.

Example 2.7.12. Let $A = \begin{bmatrix} 1 & 3 & 3 & 2 & 4 \\ 2 & 7 & 6 & 3 & 9 \\ 1 & 2 & 3 & 3 & 3 \end{bmatrix}$. Construct a basis for $\text{Nul}(A)$.

Remember that $\text{Nul}\, A$ is the solution set to $A\boldsymbol{x} = \boldsymbol{0}$, which we now solve by row reducing the matrix A. We might recognize that this is the same matrix from Example 2.7.12. As such, we have already seen that

$$[\ A \mid \boldsymbol{0}\ ] \sim \left[\begin{array}{ccccc|c} \boxed{1} & 3 & 3 & 2 & 4 & 0 \\ 0 & \boxed{1} & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

From echelon form, we can identify the pivot and non-pivot columns. Thus, we see that $\text{nullity}(A) = 3$. To find a specific basis, we are advantaged to continue to reduce the matrix to row-reduced echelon form, as shown below:

$$[\ A \mid \boldsymbol{0}\ ] \sim \left[\begin{array}{ccccc|c} \boxed{1} & 0 & 3 & 5 & 1 & 0 \\ 0 & \boxed{1} & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

As we have seen many times, this augmented matrix corresponds to a linear system, namely

$$\begin{cases} x_1 + 3x_3 + 5x_4 + x_5 = 0 \\ x_2 - x_4 + x_5 = 0. \end{cases}$$

In usual tradition, we solve for the two dependent variables, x_1, x_2, in terms of the three free variables, x_3, x_4, x_5. From this we see

$$\begin{cases} x_1 = -3x_3 - 5x_4 - x_5 \\ x_2 = x_4 - x_5, \end{cases}$$

which provides the parametric equations to the flat which is the solution set to $A\boldsymbol{x} = \boldsymbol{0}$. Converting these parametric equations into a single vector equation, we see

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -3x_3 - 5x_4 - x_5 \\ x_4 - x_5 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = x_3 \begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = x_3\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w}. \quad (2.7.1)$$

Then $\text{Nul}(A) = \text{Span}\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$. Thus, $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a spanning set for $\text{Nul}(A)$. On the other hand, the construction of $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ guarantees that they are linearly independent since

$$x_3\boldsymbol{u} + x_4\boldsymbol{v} + x_5\boldsymbol{w} = \boldsymbol{0}$$

implies that x_3, x_4, and x_5 are all zero. To see this, consider the 3rd, 4th, and 5th components of $\boldsymbol{x}$ in (2.7.1). As $\boldsymbol{u}$ contains a 1 in the 3rd component and $\boldsymbol{v}$ and $\boldsymbol{w}$ both contain 0, any combination of $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ to form $\boldsymbol{x}$ must have the coefficient of $\boldsymbol{u}$ be x_3. The same is also true that the coefficients of $\boldsymbol{v}$ and $\boldsymbol{w}$ must be x_4 and x_5, respectively. Hence, if $\boldsymbol{x}$ were $\boldsymbol{0}$, then $x_3 = x_4 = x_5 = 0$, proving the claim. Therefore, $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a basis for $\text{Nul}(A)$. ∎

Let us attempt to abbreviate the process given in the previous example (also in Example 2.6.5). First, compute the RREF of the matrix A. The basis of (A) will have exactly nullity(A) many vectors, each vector corresponding to a free variable/non-pivot column. In those entries indexed by a free variable in the vectors we are building, place a 1 or 0 according to whether the vector corresponds to the free variable from that index. For example, in the previous example, the free variables were x_3, x_4, and x_5. Thus, we start building three vector templates, in this same order, of the form

$$\begin{bmatrix} * \\ * \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ * \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ * \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

where $*$ designates a yet unspecified scalar. To fill in the *'s, notice that there indices correspond to the dependent variables/pivot columns. These pivots also correspond to pivot rows in the RREF of A, which each contain a scalar in non-pivot columns corresponding to free variables. From the first row, record the inverse of the scalar you see in the corresponding columns into the vectors we are building. For example, the first pivot row tells us to fill in the *'s as

$$\begin{bmatrix} -3 \\ * \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ * \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ * \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

We have to switch the signs of the scalars as this is a consequence of moving the variables from the left-hand side to the right-hand side of the parametric equations. Applying this same principle to Row 2, we get the same basis for (A), namely,

$$\begin{bmatrix} -3 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Example 2.7.13. Let

$$A = \begin{bmatrix} 1 & -3 & 4 & 2 & -2 \\ -2 & 6 & -8 & 5 & 3 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & -3 & 4 & 0 & -1/9 \\ 0 & 0 & 0 & \boxed{1} & -1/9 \end{bmatrix}.$$

Find a basis for $\text{Col}(A)$ and $\text{Nul}(A)$.

Since we have the RREF of A already, we can very quickly determine that $\text{Col}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \end{bmatrix}\right\}$, as a basis is formed by the 1st and 4th columns of A. Likewise, since the 2nd, 3rd, and 5th columns of A

are non-pivot columns, we see that a template for a basis of $\text{Nul}(A)$ can be begun as

$$\begin{bmatrix} * \\ 1 \\ 0 \\ * \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ 0 \\ 1 \\ * \\ 0 \end{bmatrix}, \begin{bmatrix} * \\ 0 \\ 0 \\ * \\ 1 \end{bmatrix}.$$

Reading off the scalars in the first row tells us

$$\begin{bmatrix} 3 \\ 1 \\ 0 \\ * \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 1 \\ * \\ 0 \end{bmatrix}, \begin{bmatrix} 1/9 \\ 0 \\ 0 \\ * \\ 1 \end{bmatrix}.$$

Reading off the scalars in the second row gives us a basis for $\text{Nul}(A)$, namely

$$\begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1/9 \\ 0 \\ 0 \\ 1/9 \\ 1 \end{bmatrix}.$$

Now, one can interchange a vector in a spanning set with any non-zero multiple of that same vector without changing the span. Likewise, linear independence is not affected by interchanging a non-zero multiple of a vector. Thus, if we do not want the fractions in the last spanning vector $(1/9, 0, 0, 1/9, 1) = 1/9(1, 0, 0, 1, 9)$, then we may substitute it with simply $(1, 0, 0, 1, 9)$. Therefore,

$$\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 9 \end{bmatrix}\right\}. \qquad \blacksquare$$

Exercises

(Go to Solutions)

For Exercises 1-1, determine with the statement is true or false. If false, correct the statement so that it is true.

1. If A is an $m \times n$ matrix, then $\text{rank}(A) + \text{nullity}(A) = m$.

For Exercises 2-3, let S be the provided set of vectors. Let $W = \text{Span}(S)$. Prune the spanning set S down to find a basis for W. Answers may vary.

2. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}, \begin{bmatrix} 21 \\ 16 \\ 43 \end{bmatrix}, \begin{bmatrix} 6 \\ 12 \\ 42 \end{bmatrix} \right\}$

3. $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

For Exercises 4-19, let A be the provided matrix. Find a basis for $\text{Col}(A)$ consisting of column vectors of A and a basis for $\text{Nul}(A)$. Find the rank and nullity of A. Answers may vary.

4. $\begin{bmatrix} 1 & 15 & 8 \\ 0 & 9 & 6 \\ 0 & 0 & 2 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 1 & 2 \\ 2 & 3 & 5 \\ 4 & 2 & 4 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 3 & 4 & 2 \\ 3 & 1 & 3 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 5 & 4 & 2 & 1 \\ 4 & 4 & 2 & 8 \\ 2 & 5 & 6 & 3 \end{bmatrix}$

♠ 8. $\begin{bmatrix} 8 & -3 & -13 & 15 & 17 \\ -2 & 1 & 3 & -3 & -3 \end{bmatrix}$

♠ 9. $\begin{bmatrix} -1 & 2 & 1 & 0 \\ 7 & -14 & -7 & -8 \\ -3 & 6 & 3 & 2 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 2 & 5 & 9 \\ 3 & 4 & 7 & 1 \end{bmatrix}$

11. $\begin{bmatrix} 6 & 5 & 4 & 3 & 2 & 1 \\ 5 & 4 & 3 & 2 & 1 & 6 \\ 2 & 4 & 6 & 8 & 2 & 4 \\ 0 & 1 & 0 & 1 & 2 & 6 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \\ 3 & 6 & 9 & 12 & 15 \\ 2 & -2 & 4 & 4 & 8 \end{bmatrix}$

13. $\begin{bmatrix} 1+i & 4-i \\ 2+4i & 4-2i \\ 6-10i & 3+i \\ -8-2i & 12+10i \end{bmatrix}$

♠ 14. $\begin{bmatrix} 1+2i & 2-i & -1 & 0 & 1+i & -3+3i \\ 2+4i & 4-2i & 1+4i & -2 & 2+5i & -2+10i \\ 3-i & -1-3i & 2 & i & 3-i & 6+3i \end{bmatrix}$

♠ 15. $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix} \pmod 2$

16. $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \pmod 2$

♠ 17. $\begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 0 & 3 \\ 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix} \pmod 5$

18. $\begin{bmatrix} 6 & 2 & 1 & 0 \\ 1 & 5 & 6 & 6 \\ 3 & 6 & 3 & 2 \end{bmatrix} \pmod 7$

19. $\begin{bmatrix} 1 & 2 & 6 & 8 & 9 & 3 \\ 1 & 3 & 0 & 0 & 1 & 2 \\ 1 & 4 & 2 & 8 & 7 & 9 \end{bmatrix} \pmod{11}$

20. We say that a vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ is **strongly positive** if $v_i > 0$ for all i. Let $W \leq \mathbb{R}^n$ be a subspace that contains a strongly positive vector. Prove that W has a positive consisting of only strongly positive vectors.

†There are no dependence relations on the vectors in $\emptyset$!

††There are four fundamental subspaces associated to a matrix in total. Two other fundamental subspaces, namely the Row Space and Left Null Space, will be defined later.

‡These re-definitions of rank and nullity at first view may appear to be in conflict with the Definition 1.5.5 and Definition 2.6.3, which were defined by counting pivots. Theorem 2.7.8 and the technique developed after Theorem 2.7.11 show that these notions, in fact, coincide.

"If you want to get each individual's honest opinion, you don't want that opinion to be influenced by others who are present, much less allow a group to coordinate what they are going to say." – Thomas Sowell

2.8 Coordinates

Let V be a vector space with basis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Let $\boldsymbol{x} \in V$. Since $\mathcal{B}$ is a spanning set for V, there exists scalars $c_1, c_2, \ldots, c_n \in F$ such that

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n.$$

Suppose that $\boldsymbol{x}$ can be expressed as a linear combination of $\mathcal{B}$ in another way, say

$$\boldsymbol{x} = d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \ldots + d_n\boldsymbol{v}_n,$$

for $d_1, d_2, \ldots, d_n \in F$. Then

$$\boldsymbol{0} = \boldsymbol{x} - \boldsymbol{x} = (c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n) - (d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \ldots + d_n\boldsymbol{v}_n) = (c_1 - d_1)\boldsymbol{v}_1 + (c_2 - d_2)\boldsymbol{v}_2 + \ldots + (c_n - d_n)\boldsymbol{v}_n.$$

But $\mathcal{B}$ is linearly independent, which implies that each $c_i - d_i = 0$, that is, $c_i = d_i$. Therefore, each element of V can be expressed *uniquely* as a linear combination of $\mathcal{B}$.

Definition 2.8.1. Suppose the set $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for vector space V. For each $\boldsymbol{x} \in V$, the **coordinates** of $\boldsymbol{x}$ **relative to** the basis $\mathcal{B}$ are the unique coefficients $c_1, \ldots, c_n \in F$ such that

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n.$$

The vector $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \in F^n$ is called the **coordinate vector** of $\boldsymbol{x}$ **relative to** $\mathcal{B}$ or the $\mathcal{B}$**-coordinate vector** of $\boldsymbol{x}$.

Example 2.8.2. Let $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$, and $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$. Then $\mathcal{B}$ is a basis for $V = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ because $\mathcal{B}$ is linearly independent. Determine if $\boldsymbol{x} \in V$, and if it is, find the coordinate vector of $\boldsymbol{x}$ relative to $\mathcal{B}$.

If $\boldsymbol{x} \in V$, then there exists $c_1, c_2 \in \mathbb{R}$ such that $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 = \boldsymbol{x}$. Then, we row reduce the corresponding augmented matrix, below:

$$\left[\begin{array}{cc|c} 1 & 5 & 3 \\ 0 & 2 & 2 \\ 1 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right].$$

From this we see that $\boldsymbol{x} \in V$ and

$$\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}.$$

Therefore, $[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$. The basis $\mathcal{B}$ determines a "coordinate system" for the plane spanned by $\boldsymbol{v}_1$, $\boldsymbol{v}_2$. ∎

In the previous example, even though the vectors in V are vectors in $\mathbb{R}^3$, they are completely determined by their coordinate vectors, which belong to $\mathbb{R}^2$. Thus, there is a natural identification between the vectors of W and the vectors in $\mathbb{R}^2$, namely $\boldsymbol{x} \mapsto [\boldsymbol{x}]_{\mathcal{B}}$. This mapping is a one-to-one, onto linear transformation. Essentially, this means that the two spaces *look* the same. In this example, V and $\mathbb{R}^2$ are geometrically the same as they are both planes. We even write that $V \cong \mathbb{R}^2$, and we say that W is **congruent** to (or **isomorphic** to) to $\mathbb{R}^2$.

Example 2.8.3. Consider two bases $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2\}$ for a vector space V, such that

$$\boldsymbol{c}_1 = 2\boldsymbol{b}_1 - 3\boldsymbol{b}_2 \qquad \text{and} \qquad \boldsymbol{c}_2 = -3\boldsymbol{b}_1 + 5\boldsymbol{b}_2.$$

Suppose further that

$$\boldsymbol{x} = \boldsymbol{c}_1 + 3\boldsymbol{c}_2.$$

Compute $[\boldsymbol{x}]_{\mathcal{B}}$.

We already know that

$$[\boldsymbol{c}_1]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \qquad [\boldsymbol{c}_2]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}, \quad \text{and} \quad [\boldsymbol{x}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

Thus,

$$[\boldsymbol{x}]_{\mathcal{B}} = [\boldsymbol{c}_1 + 3\boldsymbol{c}_2]_{\mathcal{B}} = [\boldsymbol{c}_1]_{\mathcal{B}} + 3[\boldsymbol{c}_2]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} + 3\begin{bmatrix} -3 \\ 5 \end{bmatrix} = \begin{bmatrix} -7 \\ 12 \end{bmatrix}.$$

On the other hand, the vector equation

$$[\boldsymbol{x}]_{\mathcal{B}} = [\boldsymbol{c}_1]_{\mathcal{B}} + 3[\boldsymbol{c}_2]_{\mathcal{B}}$$

can be rewritten as a matrix equation

$$[\boldsymbol{x}]_{\mathcal{B}} = \begin{bmatrix} [\boldsymbol{c}_1]_{\mathcal{B}} & [\boldsymbol{c}_2]_{\mathcal{B}} \end{bmatrix} [\boldsymbol{x}]_{\mathcal{C}} \quad \Rightarrow \quad \begin{bmatrix} -7 \\ 12 \end{bmatrix} = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$ ■

Theorem 2.8.4. *Let $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n\}$ be bases of a vector space V. Then there is a unique $n \times n$ matrix $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$, called the **change-of-basis matrix** to $\mathcal{B}$ from $\mathcal{C}$ such that*

$$[\boldsymbol{x}]_{\mathcal{B}} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[\boldsymbol{x}]_{\mathcal{C}}.$$

The columns of $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ are the $\mathcal{B}$-coordinate vectors of the elements of $\mathcal{C}$, that is,

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} [\boldsymbol{c}_1]_{\mathcal{B}} & [\boldsymbol{c}_2]_{\mathcal{B}} & \cdots & [\boldsymbol{c}_n]_{\mathcal{B}} \end{bmatrix}.$$

Multiplication by $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ converts $\mathcal{C}$-coordinate vectors to $\mathcal{B}$-coordinate vectors. To change coordinates between two bases, we need the coordinates of the old basis in terms of the new basis.

Example 2.8.5. Let $\boldsymbol{b}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$, $\boldsymbol{c}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\boldsymbol{c}_2 = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2\}$. Then both $\mathcal{B}$ and $\mathcal{C}$ are bases of $\mathbb{R}^2$. Find the change-of-basis matrix $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$.

We need to compute the coordinate vectors $[\boldsymbol{c}_1]_{\mathcal{B}}$ and $[\boldsymbol{c}_2]_{\mathcal{B}}$. We compute the first coordinate vector by solving the linear system

$$\left[\begin{array}{cc|c} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{c}_1 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 4 & 1 \\ 2 & 3 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -5 \\ 0 & 1 & 4 \end{array}\right] \quad \Rightarrow \quad [\boldsymbol{c}_1]_{\mathcal{B}} = \begin{bmatrix} -5 \\ 4 \end{bmatrix}.$$

Using the same row operations, we see that

$$\left[\begin{array}{cc|c} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \boldsymbol{c}_2 \end{array}\right] = \left[\begin{array}{cc|c} 3 & 4 & 5 \\ 2 & 3 & 1 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 11 \\ 0 & 1 & -7 \end{array}\right] \quad \Rightarrow \quad [\boldsymbol{c}_2]_{\mathcal{B}} = \begin{bmatrix} 11 \\ -7 \end{bmatrix}.$$

This gives

$$\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} -5 & 11 \\ 4 & -7 \end{bmatrix}. \qquad \blacksquare$$

Let $\mathcal{B}$ and $\mathcal{C}$ be two bases of a vector space V. Then mimicking the previous example, we can see that

$$\left[\begin{array}{c|c} \mathcal{B} & \mathcal{C} \end{array}\right] \sim \left[\begin{array}{c|c} \mathcal{E} & \underset{\mathcal{B}\leftarrow\mathcal{C}}{P} \end{array}\right], \tag{2.8.1}$$

where $\mathcal{E}$ denotes the standard basis for F^n. We should mention that the echelon form of $\mathcal{B}$ might contain rows of zeros. Thus, the left-hand side of this row-reduced echelon form really should be of the form $\begin{bmatrix} \mathcal{E} \\ 0 \end{bmatrix}$, where here 0 denotes some matrix consisting of only zeros. If, in fact, $\mathcal{B}$ and $\mathcal{C}$ really do not span the same vector space, then there would instead be a nonzero entry on the right side of one of these rows of zeros, indicating an inconsistent linear system. When the linear system is consistent though, then $\mathcal{B}$ and $\mathcal{C}$ truly were bases for the same span and on the right-hand side of each of these rows of zeros will be accompanying rows of zeros. Thus, (2.8.1) more properly should be $\left[\begin{array}{c|c} \mathcal{B} & \mathcal{C} \end{array}\right] \sim \left[\begin{array}{c|c} \mathcal{E} & \underset{\mathcal{B}\leftarrow\mathcal{C}}{\mathcal{P}} \\ 0 & 0 \end{array}\right]$.

Example 2.8.6. Let $\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \\ -3 \\ 0 \end{bmatrix}$, $\boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 4 \end{bmatrix}$, $\boldsymbol{b}_3 = \begin{bmatrix} 3 \\ 0 \\ 0 \\ -2 \end{bmatrix}$, $\boldsymbol{c}_1 = \begin{bmatrix} 6 \\ 3 \\ -21 \\ 26 \end{bmatrix}$, $\boldsymbol{c}_2 = \begin{bmatrix} 15 \\ 5 \\ -23 \\ 12 \end{bmatrix}$, $\boldsymbol{c}_3 = \begin{bmatrix} 3 \\ 2 \\ -8 \\ 4 \end{bmatrix}$, $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3\}$. Then both $\mathcal{B}$ and $\mathcal{C}$ are bases for the same subspace of $\mathbb{R}^4$.

(a) Find the change-of-basis matrix to $\mathcal{B}$ from $\mathcal{C}$.

$$\left[\begin{array}{c|c} \mathcal{B} & \mathcal{C} \end{array}\right] = \left[\begin{array}{ccc|ccc} 1 & 1 & 3 & 6 & 15 & 3 \\ 1 & 0 & 0 & 3 & 5 & 2 \\ -3 & -2 & 0 & -21 & -23 & -8 \\ 0 & 4 & -2 & 26 & 12 & 4 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & 5 & 2 \\ 0 & 1 & 0 & 6 & 4 & 1 \\ 0 & 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P} = \begin{bmatrix} 3 & 5 & 2 \\ 6 & 4 & 1 \\ -1 & 2 & 0 \end{bmatrix}$.

(b) If $[\boldsymbol{x}]_{\mathcal{C}} = \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix}$, then compute $[\boldsymbol{x}]_{\mathcal{B}}$.

$$[\boldsymbol{x}]_{\mathcal{B}} = \underset{\mathcal{B}\leftarrow\mathcal{C}}{P}[\boldsymbol{x}]_{\mathcal{C}} = \begin{bmatrix} 3 & 5 & 2 \\ 6 & 4 & 1 \\ -1 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -8 \end{bmatrix}.$$

Note that

$$\boldsymbol{x} = -\boldsymbol{b}_1 + 4\boldsymbol{b}_2 - 8\boldsymbol{b}_3 = 2\boldsymbol{c}_1 - 3\boldsymbol{c}_2 + 4\boldsymbol{c}_3 = \begin{bmatrix} -21 \\ -1 \\ -5 \\ 32 \end{bmatrix}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-3, find the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$ given the basis $\mathcal{B}$, where $\boldsymbol{x}$ is the vector provided first and $\mathcal{B}$ is the set of vectors provided second.

♠ 1. $\begin{bmatrix} 4 \\ -25 \\ 6 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} \right\}$

♠ 2. $\begin{bmatrix} -1 \\ 22 \\ 3 \\ 2 \end{bmatrix}, \left\{ \begin{bmatrix} 2 \\ 5 \\ -3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 4 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

♠ 3. $\begin{bmatrix} -2-3i \\ -5+6i \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} 1+2i \\ 2-3i \end{bmatrix} \right\}$

For Exercises 4-7, given the coordinate vector $[\boldsymbol{x}]_{\mathcal{C}}$ and the change-of-basis matrix $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$, compute the coordinate vector $[\boldsymbol{x}]_{\mathcal{B}}$, where $[\boldsymbol{x}]_{\mathcal{C}}$ is the vector provided first and $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ is the matrix provided second.

♠ 4. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 & -2 & 1 \\ 3 & 0 & 2 \\ 5 & 6 & -2 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 & -2 & -1 & 0 \\ 3 & 2 & 2 & -1 \\ 0 & 0 & 8 & 2 \\ -1 & -2 & 2 & 1 \end{bmatrix}$

♠ 6. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 6 & 7 & 2 \\ 1 & 3 & 4 \\ 5 & 5 & 1 \end{bmatrix} \pmod{11}$

7. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5-6i & -3-i & 8+7i \\ 3+i & 2+i & 1+i \\ 4+2i & -9-6i & 3-i \end{bmatrix}$

For Exercises 8-13, compute the change-of-basis matrix $\underset{\mathcal{B}\leftarrow\mathcal{C}}{P}$ for the bases $\mathcal{B}$ and $\mathcal{C}$.

♠ 8. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ -5 \end{bmatrix} \right\}, \mathcal{C} = \left\{ \begin{bmatrix} 4 \\ 8 \\ 5 \end{bmatrix}, \begin{bmatrix} -5 \\ -12 \\ 20 \end{bmatrix}, \begin{bmatrix} -12 \\ -28 \\ 27 \end{bmatrix} \right\}$

♠ 9. $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 0 \\ 7 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ -2 \\ 5 \end{bmatrix} \right\}, \mathcal{C} = \left\{ \begin{bmatrix} -52 \\ 1 \\ -1 \\ 72 \end{bmatrix}, \begin{bmatrix} -20 \\ 2 \\ 2 \\ 88 \end{bmatrix}, \begin{bmatrix} -6 \\ -3 \\ -9 \\ 29 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 3 \\ 52 \end{bmatrix} \right\}$

10. $\mathcal{B} = \left\{ \begin{bmatrix} 2 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} \right\}, \mathcal{C} = \left\{ \begin{bmatrix} -15 \\ 35 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -15 \\ 25 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 4 \\ 10 \\ -6 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -8 \\ 4 \\ 4 \end{bmatrix} \right\}$

♠ 11. $\mathcal{B} = \left\{ \begin{bmatrix} 1+2i \\ 2-3i \end{bmatrix}, \begin{bmatrix} 1-i \\ 1+4i \end{bmatrix} \right\}, \mathcal{C} = \left\{ \begin{bmatrix} 4i \\ 1 \end{bmatrix}, \begin{bmatrix} 7+6i \\ i \end{bmatrix} \right\}$

12. $\mathcal{B} \equiv \left\{ \begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} \right\}, \mathcal{C} \equiv \left\{ \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 4 \end{bmatrix} \right\} \pmod 5$

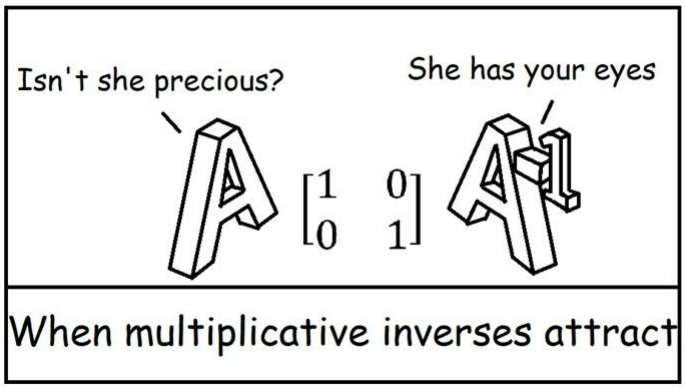
13. $\mathcal{B} \equiv \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}, \mathcal{C} \equiv \left\{ \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\} \pmod 5$

Chapter 3

The Algebra and Geometry of Matrices

Up to this point, we have used matrices for one purpose, to encode information about linear systems. We first introduced the augmented matrix in Section 1.5 to do exactly that. In Chapter 2 we started using matrices to represent a set of vectors in a slightly more compact way. This happened in Section 2.2, where we introduce the matrix-vector product so that the matrix equation (2.2.1) could encode the vector equation (2.1.1). This allowed us to define linear transformations using matrices.

But we also introduced the the column and null spaces of a matrix. These vector spaces came about essentially by solving problems about spanning and linear independence of vectors but we attributed the spaces to the matrix, not a collection of vectors. This was our first inkling that matrices deserve to be studied in their own right, not just as tools to better understand vectors. In fact, we will turn this paradigm on its head. Viewing column vectors just as $n \times 1$ matrices, all we have studied about vectors can actually be viewed as a subset of the theory of matrices. We can add and scale matrices just like column vectors. For this reason, we can actually view matrices as just beefier versions of column vectors, which can lead to discussions about vector spaces of matrices. But matrices can transcend the vector operations as we will introduce a general matrix multiplication, which will be the foundation for this chapter on matrices.



"The thing that scares us the most is when familiar things operate in unfamiliar ways." – Noah Hawley

3.1 Matrix Operations

We say that two $m \times n$ matrices $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ and $B = \begin{bmatrix} b_{ij} \end{bmatrix}$ are **equal** if $a_{ij} = b_{ij}$ for all i and j (for all $a_{ij}, b_{ij} \in F$). We can **add** matrices term-wise, that is,

$$A + B = \begin{bmatrix} a_{ij} + b_{ij} \end{bmatrix}.$$

Finally, we can **multiply** matrices by a **scalar** term-wise, that is,

$$cA = \begin{bmatrix} ca_{ij} \end{bmatrix} \qquad \text{for all } c \in F.$$

Example 3.1.1. Let $A = \begin{bmatrix} 3 & 9 & 1 \\ -2 & 4 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 5 & 6 \\ 3 & 1 & 1 \end{bmatrix}$, and $C = \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix}$.

Then

$$A + B = \begin{bmatrix} 3+0 & 9+5 & 1+6 \\ -2+3 & 4+1 & 6+1 \end{bmatrix} = \begin{bmatrix} 3 & 14 & 7 \\ 1 & 5 & 7 \end{bmatrix}.$$

Now, $A + C$ is not possible since the matrices have different sizes.

Next,

$$2B = 2\begin{bmatrix} 0 & 5 & 6 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 10 & 12 \\ 6 & 2 & 2 \end{bmatrix},$$

and

$$A - 2B = \begin{bmatrix} 3 & 9 & 1 \\ -2 & 4 & 6 \end{bmatrix} - \begin{bmatrix} 0 & 10 & 12 \\ 6 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 3 & -1 & -11 \\ -8 & 2 & 4 \end{bmatrix}.$$ ∎

Because we can add and scale matrices, we can actually view matrices as vector themselves. Let F be a field. Then $F^{m\times n}$ will denote the set of $m \times n$ matrices with entries from F, which is a vector space. This means that addition of matrices and multiplication of scalars follow the eight algebraic properties listed in Definition 1.3.1. When working with matrices, we will let 0 denote an $m \times n$ matrix with all entries equal to zero. This is called the **zero matrix**.

Let E_{ij} be the matrix whose entry in the (i, j)th entry is a one and all other entries are zero. Then $\mathcal{E} = \{E_{1,1}, E_{1,2}, \ldots, E_{1,n}, E_{2,1}, E_{2,2}, \ldots, E_{2,n}, \ldots, E_{m,1}, E_{m,2}, \ldots, E_{m,n}\}$ forms the **standard basis** of $F^{m\times n}$. The matrices $E_{i,j}$ are often called the unit matrices. Using coordinate vectors, we see that $F^{m\times n} \cong F^{mn}$, since $|\mathcal{E}| = mn$.

Definition 3.1.2. The **diagonal entries** in an $m \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ are the entries a_{11}, a_{22}, a_{33}, …, that is, the entries a_{ii}, and they form the **main diagonal**.

If A is an $n \times n$ matrix, then we say that A is a **square matrix**.

Let I_n denote the $n \times n$ matrix whose entries are 1's across the diagonal and 0's everywhere else. This matrix is known as the **identity matrix**.

Definition 3.1.3. If A is an $m \times n$ matrix and B is an $n \times p$ matrix with column vectors

$$B = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \dots & \boldsymbol{b}_p \end{bmatrix},$$

then the **matrix product**

$$AB = A \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \dots & \boldsymbol{b}_p \end{bmatrix} = \begin{bmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \dots & A\boldsymbol{b}_p \end{bmatrix}.$$

The matrix AB is a $m \times p$ matrix.

If A is an $n \times n$ matrix, then $A^k = \underbrace{A \cdots A}_{k}$. We let $A^0 = I_n$.

Matrix multiplication can also be defined with the seemingly complicated formula,

$$AB = \begin{bmatrix} \sum_{k=1}^{n} a_{ik}b_{kj} \end{bmatrix} = \begin{bmatrix} a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{in}b_{nj} \end{bmatrix} \qquad \text{for } A = \begin{bmatrix} a_{ik} \end{bmatrix} \text{ and } B = \begin{bmatrix} b_{kj} \end{bmatrix},$$

which is none other than the "finger-multiplication" we learned with the matrix-vector product.

Example 3.1.4. Compute AB with the given matrices $A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 4 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 9 & -5 & -3 \\ 3 & 9 & 1 \\ -2 & 4 & 6 \end{bmatrix}$.

$$AB = \begin{bmatrix} A\begin{bmatrix} 9 \\ 3 \\ -2 \end{bmatrix} & A\begin{bmatrix} -5 \\ 9 \\ 4 \end{bmatrix} & A\begin{bmatrix} -3 \\ 1 \\ 6 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 2(9)+1(3)-1(-2) & 2(-5)+1(9)-1(4) & 2(-3)+1(1)-1(6) \\ 0(9)+4(3)-2(-2) & 0(-5)+4(9)-2(4) & 0(-3)+4(1)-2(6) \end{bmatrix}$$

$$= \begin{bmatrix} 18+3+2 & -10+9-4 & -6+1-6 \\ 0+12+4 & 0+36-8 & 0+4-12 \end{bmatrix} = \begin{bmatrix} 23 & -5 & -11 \\ 16 & 28 & -8 \end{bmatrix},$$

which is a 2×3 matrix. ∎

Definition 3.1.5. If A is a square matrix, say $n \times n$, and if

$$p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_mx^m$$

is a degree m polynomial, then we define the $n \times n$ matrix $p(A)$ to be

$$p(A) = a_0I_n + a_1A + a_2A^2 + \ldots + a_mA^m.$$

An expression of this form is called a **matrix polynomial** in A.

Example 3.1.6. Find $p(A)$ for $p(x) = x^2 + 3x + 2$ and $A = \begin{bmatrix} 4 & 2 \\ 0 & 3 \end{bmatrix}$ over $\mathbb{Z}_5$.

$$p(A) \equiv A^2 + 3A + 2I \equiv \begin{bmatrix} 4 & 2 \\ 0 & 3 \end{bmatrix}^2 + 3\begin{bmatrix} 4 & 2 \\ 0 & 3 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\equiv \begin{bmatrix} 1 & 4 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 0 & 4 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \equiv \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \blacksquare$$

Definition 3.1.7. Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ be an $m \times n$ matrix. Then the **transpose** of A, denoted $A^\top$, is the $n \times m$ matrix given by $A^\top = \begin{bmatrix} a_{ji} \end{bmatrix}$, that is, the matrix whose columns are formed from the corresponding rows of A.

Example 3.1.8. Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \; B = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 7 \end{bmatrix}, \text{ and } \; C = \begin{bmatrix} 2 & -3 \\ 0 & 1 \end{bmatrix}.$$

Then

$$A^\top = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad B^\top = \begin{bmatrix} 1 & 3 \\ 1 & 5 \\ 1 & 7 \end{bmatrix}, \text{ and } C^\top = \begin{bmatrix} 2 & 0 \\ -3 & 1 \end{bmatrix}.$$

$\blacksquare$

In the case of $\mathbb{C}$-matrices, an alternative to transposes is preferred for reasons that will be explained in Chapter 4.

Definition 3.1.9. Let A be an $m \times n$ complex matrix. Then we define $A^* = (\overline{A})^\top$, which is called the **conjugate transpose**. This replaces the role of transposes in complex space.

When discussing transposes of matrices, whenever the matrices are complex, the conjugate transpose should ALWAYS be used instead of the standard transpose.

Example 3.1.10. Let $A = \begin{bmatrix} 1-2i & 3+5i & 6 \\ -2i & 0 & i \end{bmatrix}$ and $B = \begin{bmatrix} 2-3i & 0 & 2i \\ -i & 4 & 1+2i \\ 0 & 2-2i & 6 \end{bmatrix}$.

Note that

$$A^* = \begin{bmatrix} 1+2i & 2i \\ 3-5i & 0 \\ 6 & -i \end{bmatrix}, \quad B^* = \begin{bmatrix} 2+3i & i & 0 \\ 0 & 4 & 2+2i \\ -2i & 1-2i & 6 \end{bmatrix}. \quad \blacksquare$$

Definition 3.1.11. If A is a square matrix, then the **trace** of A, denoted $\mathrm{tr}(A)$, is the sum of the diagonal entries of A.

Example 3.1.12. Let $A = \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 9 & 3 & -2 & 3 \\ 28 & 9 & 4 & 0 \\ 16 & 1 & 6 & 3 \\ 2 & 5 & 2 & 1 \end{bmatrix}$. Then

$$\mathrm{tr}(A) = 2 + 4 = \boxed{6}, \qquad \mathrm{tr}(B) = 9 + 9 + 6 + 1 = \boxed{25}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-6, determine with the statement is true or false. If false, correct the statement so that it is true.

1. We say that two $m \times n$ matrices $A = [a_{ij}]$ and $B = [b_{ij}]$ are equal if $a_{ij} = b_{ij}$ for all i and j.
2. If A is a square matrix, then the trace of A is the sum of the column entries of A.
3. We can multiply matrices by a scalar term-wise.
4. The identity matrix is an $n \times n$ matrix whose entries are 1's across the main diagonal and zeros everywhere else.
5. If A is an $n \times n$ matrix, then we say that A is a square matrix.
6. The transpose of a matrix A is a matrix whose columns are formed from the corresponding rows of A.
7. Write the matrix $E_{1,3} \in F^{3\times 3}$.

For Exercises 8-10, using the matrices listed below, explain why the operation is not possible.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 3 & 6 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 7 & 5 \\ 1 & 5 & 4 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 6 & 4 \\ 5 & 9 & 2 \\ 9 & 3 & 3 \end{bmatrix}.$$

8. CA
9. ABD
10. $\text{tr}(C)$

For Exercises 11-20, using the matrices listed below, perform the matrix calculation:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & -3 & 0 \\ 1 & -2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 5 & 6 \\ -5 & -5 & 2 \\ 2 & 0 & -3 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 0 & -1 & 2 \\ 1 & 7 & 8 & -3 \\ 3 & 3 & 2 & -4 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 3 & -2 \\ -5 & 0 & 7 \\ 0 & -2 & 4 \\ 1 & 2 & 3 \end{bmatrix}.$$

♠ 11. $2A + B^\top$
♠ 12. $4A - 3B$
♠ 13. AC
♠ 14. $A^\top$
♠ 15. $C^\top$
♠ 16. $D^\top$
♠ 17. $\text{tr}(A)$
♠ 18. $\text{tr}(B)$
♠ 19. $DB^\top$
♠ 20. $A^2 + 2A - 5I_3$

For Exercises 21-24, using the matrices listed below, perform the matrix calculation over the field $\mathbb{R}$:

$$A = \begin{bmatrix} -1 & 0 & 2 \\ 1 & 3 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 1 \\ 2 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} -2 & 0 \\ 3 & -2 \end{bmatrix}.$$

21. $2A^\top + B$
22. AB
23. BC
24. ABC

For Exercises 25-33, using the matrices listed below, perform the matrix calculation over the field $\mathbb{C}$:

$$A = \begin{bmatrix} 1+2i & 1-3i \\ 5 & 1-i \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 3-4i \\ 3i & 1+i \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 0 \\ i & 2i \\ 3+i & 1+3i \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & i & 1-2i \\ 5 & 4+i & 0 & -4 \end{bmatrix}.$$

♠ 25. $(1+i)A-3B$ ♠ 26. CA ♠ 27. A^* ♠ 28. C^* ♠ 29. D^*

♠ 30. $\mathrm{tr}(A)$ ♠ 31. $\mathrm{tr}(B)$ ♠ 32. $(BD)^*$ ♠ 33. A^2+I_2

For Exercises 34-42, using the matrices listed below, perform the matrix calculation over the field $\mathbb{Z}_5$:

$$A \equiv \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 1 & 2 & 2 & 1 \\ 4 & 4 & 3 & 2 \end{bmatrix}, \quad B \equiv \begin{bmatrix} 0 & 1 & 1 & 2 \\ 2 & 2 & 3 & 0 \\ 1 & 2 & 3 & 2 \\ 4 & 3 & 4 & 2 \end{bmatrix}, \quad C \equiv \begin{bmatrix} 3 & 0 & 4 & 2 \\ 1 & 2 & 3 & 2 \\ 3 & 3 & 2 & 1 \end{bmatrix}, \quad D \equiv \begin{bmatrix} 2 & 3 & 3 \\ 0 & 0 & 2 \\ 0 & 3 & 4 \\ 1 & 2 & 3 \end{bmatrix} \pmod 5.$$

♠ 34. $2A-3B$ ♠ 35. CA ♠ 36. $A^\top$ ♠ 37. $C^\top$ ♠ 38. $D^\top$

♠ 39. $\mathrm{tr}(A)$ ♠ 40. $\mathrm{tr}(B)$ ♠ 41. $D^\top B$ ♠ 42. A^2+A-I_4

For Exercises 43-45, for the matrix A provided, find $\mathrm{tr}(A)$, $A^\top$, and $\mathrm{tr}(A^\top)$.

43. $\begin{bmatrix} 8 & -7 & 2 \\ 0 & -2 & 4 \\ 3 & 2 & 1 \end{bmatrix}$

44. $\begin{bmatrix} 1/8 & -7/8 & 2/5 & -12/13 \\ 9/8 & -1/2 & 1/4 & 1/4 \\ 1 & -12/15 & 2 & 1/5 \\ 1/4 & 8/15 & -3/8 & 3/8 \end{bmatrix}$

45. $\begin{bmatrix} 5 & 1 & 6 & 2 \\ 3 & 0 & 3 & 4 \\ 5 & 2 & 1 & 1 \\ 2 & 3 & 0 & 0 \end{bmatrix} \pmod 7$

For Exercises 46-46, for the matrix A provided, find $\mathrm{tr}(A)$, A^*, and $\mathrm{tr}(A^*)$.

46. $A = \begin{bmatrix} 1-3i & -5-5i & 2+2i \\ 2-4i & 8i & 6+i \\ -1+2i & 3-6i & 2+4i \end{bmatrix}$

"When you put your hand to the plow, you can't put it down until you get to the end of the row."
– Alice Paul

3.2 Matrix Properties

We mentioned in the previous section that $F^{m\times n}$, meaning that matrix addition and scalar multiplication satisfy the eight properties listed in Definition 1.3.1. Transposes and traces also follow very nice algebraic properties.

Theorem 3.2.1. *Let A and B be matrices with sizes such that the indicated sums and products are defined. Let $c \in F$.*

(i) $(A+B)^\top = A^\top + B^\top$

(ii) $(cA)^\top = cA^\top$

(iii) $(A^\top)^\top = A$

(iv) $(AB)^\top = B^\top A^\top$

We see by the first two properties that transposition is a linear transformation $^\top : F^{m\times n} \to F^{n\times m}$. Note the last property states that the transpose of a product of matrices equals the product of their transposes in the reverse order! This is called the *Shoe-Sock Principle*†. As we will see momentarily (see Theorem 3.2.4), this order matters a lot.

Theorem 3.2.2. *Let A and B be $n \times n$ matrices. Let $c \in F$.*

(i) $\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$

(ii) $\operatorname{tr}(cA) = c\operatorname{tr}(A)$

(iii) $\operatorname{tr}(A^\top) = \operatorname{tr}(A)$

(iv) $\operatorname{tr}(AB) = \operatorname{tr}(BA)$

We see by the first two properties that the trace is a linear transformation $\operatorname{tr} : F^{m\times n} \to F$. It should be mentioned that the last property does NOT say that $\operatorname{tr}(AB) = \operatorname{tr}(A)\operatorname{tr}(B)$. This is, in fact, very false. Also, this property does NOT say that $AB = BA$ only that $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. As hinted to already, we will soon see that $AB \neq BA$, in general.

Unfortunately, matrix multiplication is not as well behaved as multiplication of real numbers which we are used to. Many of the typical algebraic properties do hold.

Theorem 3.2.3. *Let A, B, and C be matrices with sizes such that the indicated sums and products are defined. Let $c \in F$.*

(i) $A(BC) = (AB)C$

(ii) $A(B+C) = AB + AC$

(iii) $(A+B)C = AC + BC$

(iv) $c(AB) = (cA)B = A(cB)$

(v) Let A be an $m \times n$ matrix. Then

$$I_m A = A = A I_n.$$

Please be aware that certain multiplicative properties in the previous list are omitted. This is quite intentional because many algebraic properties that we take for granted with real multiplication do NOT hold for matrix multiplication.

Theorem 3.2.4. *Let A, B, and C be matrices with sizes such that the indicated sums and products are defined.*

(i) You CANNOT assume that $AB = BA$

(ii) You CANNOT assume that if $AB = AC$*, then* $B = C$

(iii) You CANNOT assume that if $AB = 0$*, then* $A = 0$ *or* $B = 0$*.*

Example 3.2.5. Let $A = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & -3 \\ 1 & 1 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 0 & -3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 3 & 6 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 0 & -3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & -9 \\ 0 & 5 \end{bmatrix}.$$

Therefore, $AB \neq BA$. ∎

Example 3.2.6. Let $A = \begin{bmatrix} 0 & 2 \\ 0 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, and $C = \begin{bmatrix} 5 & 6 \\ 3 & 4 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 0(1)+2(3) & 0(2)+2(4) \\ 0(1)-(3) & 0(2)-(4) \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ -3 & -4 \end{bmatrix} \quad \text{and} \quad AC = \begin{bmatrix} 0(5)+2(3) & 0(6)+2(4) \\ 0(5)-(3) & 0(6)-(4) \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ -3 & -4 \end{bmatrix}.$$

Therefore, $AB = AC$ but $B \neq C$. In other words, we cannot divide both sides by A. ∎

Example 3.2.7. Let $A = \begin{bmatrix} 0 & 2 \\ 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 2 \\ 0 & 0 \end{bmatrix}$. Then

$$AB = \begin{bmatrix} 0(4)+2(0) & 0(2)+2(0) \\ 0(4)-(0) & 0(2)-(0) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

which is the zero matrix. Therefore, $AB = 0$ but $A \neq 0$ not $B \neq C$. In other words, the zero product property does NOT hold for matrices. ∎

Definition 3.2.8. Let A be an $m \times n$ matrix. Define the **row space** of A, denoted $\operatorname{Row} A$, as the column space of $A^\top$, that is, $\operatorname{Row} A = \operatorname{Col} A^\top$.

The dimension of $\operatorname{Row} A$ is called the **corank**, denoted $\operatorname{corank} A = p$, where p is the number of pivots in A.

Note that the corank is not really a new quantity. Note that $\operatorname{corank}(A) = \dim(\operatorname{Row} A) =$ the number of pivot positions $= \operatorname{rank} A = \dim(\operatorname{Col} A)$.

For another way to describe the row space of A, note that each row of A contains n entries and can be identified with a vector in F^n. Then $\operatorname{Row} A$ is the span of the rows of A under this identification.

Theorem 3.2.9. *Two matrices* A *and* B *are row equivalent if and only if* $\operatorname{Row} A = \operatorname{Row} B$*.*

Theorem 3.2.10. *If U is an echelon form of the matrix A, then nonzero rows of U form a basis for the row space of A.*

Much like the column space of A, the pivot rows of A also form a basis of $\text{Row}(A)$. But unlike the column space, the pivot rows of U form a basis of $\text{Row}(A)$. This is not true for the column space, that is, the pivot columns of U do NOT form a basis for $\text{Col}(A)$. This is because row equivalent matrices need not have the same column space. Only the row space is guaranteed to be equal.

Example 3.2.11. Let $A = \begin{bmatrix} 1 & 3 & 2 & 4 & 2 \\ 2 & 6 & 4 & 8 & 4 \\ 3 & 2 & 5 & 2 & 1 \\ 4 & 2 & 5 & 1 & 0 \end{bmatrix}$. Compute a basis for $\text{Row}\, A$, $\text{Col}\, A$, and $\text{Nul}\, A$.

Since

$$A = \begin{bmatrix} 1 & 3 & 2 & 4 & 2 \\ 2 & 6 & 4 & 8 & 4 \\ 3 & 2 & 5 & 2 & 1 \\ 4 & 2 & 5 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 15/11 & 7/11 \\ 0 & 0 & 1 & 5/11 & 6/11 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Therefore,

$$\{(1,0,0,-1,-1),(0,1,0,15/11,7/11),(0,0,1,5/11,6/11)\}$$

is a basis for $\text{Row}\, A$. Hence, $\text{corank}(A) = 3$. If a basis without fractions is desired, those vectors can be scaled by 11 without adjusting the span or the linear independence. Thus,

$$\{(1,0,0,-1,-1),(0,11,0,15,7),(0,0,11,5,6)\}$$

is another basis of $\text{Row}\, A$.

We also see that

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 6 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 5 \\ 5 \end{bmatrix} \right\}$$

is a basis for $\text{Col}\, A$. Hence, $\text{rank}(A) = 3$.

Likewise,

$$\left\{ \begin{bmatrix} 11 \\ -15 \\ -5 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 11 \\ -7 \\ -6 \\ 0 \\ 1 \end{bmatrix} \right\}$$

is a basis for $\text{Nul}\, A$. Hence, $\text{nullity}(A) = 2$. ∎

Alternatively, one could find a basis for $\text{Row}\, A$ by finding a basis for $\text{Col}\, A^\top$ as we did before. This would, in fact, provide a basis consisting of actual rows of A. The method above has the advantage (other than providing simpler vectors) than we can find all these bases of the fundamental spaces of A simultaneously.

Exercises (Go to Solutions)

1. Rewrite the eight axioms of a vector space, listed in Definition 1.3.1, as properties of $m \times n$ matrices.

2. Verify Theorem 3.2.4 (i) using $A = \begin{bmatrix} 3 & 2 & 1 \\ 1 & 5 & 0 \\ 4 & 1 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 4 \\ 2 & 1 & 7 \end{bmatrix}$.

For Exercises 3-7, for the matrix A provided, find a basis for $\text{Row}(A)$ and compute $\text{corank}(A)$. Answers may vary.

♠ 3. $\begin{bmatrix} 8 & -3 & -13 & 15 & 17 \\ -2 & 1 & 3 & -3 & -3 \end{bmatrix}$

♠ 4. $\begin{bmatrix} -1 & 2 & 1 & 0 \\ 7 & -14 & -7 & -8 \\ -3 & 6 & 3 & 2 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$ (mod 2)

♠ 6. $\begin{bmatrix} 2 & 1 & 1 & 1 \\ 4 & 2 & 0 & 3 \\ 2 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 \end{bmatrix}$ (mod 5)

♠ 7. $\begin{bmatrix} 1+2i & 2-i & -1 & 0 & 1+i & -3+3i \\ 2+4i & 4-2i & 1+4i & -2 & 2+5i & -2+10i \\ 3-i & -1-3i & 2 & i & 3-i & 6+3i \end{bmatrix}$

For Exercises 8-8, let A be the matrix provided on the left. The second matrix is row equivalent to A. Find a basis for $\text{Col}(A)$ and $\text{Row}(A)$ and compute $\text{rank}(A)$ and $\text{corank}(A)$.

8. $\begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 1 & 2 & -4 & 10 & 13 & -12 \\ 1 & -1 & -1 & 1 & 1 & -3 \\ 1 & -3 & 1 & 5 & 7 & 3 \\ 1 & -2 & 0 & 0 & -5 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -3 & 7 & 9 & -9 \\ 0 & 1 & -4 & 3 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

9. Show that $\mathbb{R}^{2\times 2}$ with respect to matrix addition and matrix multiplication is NOT a field.

10. Show that if F is any field and n a positive integer, then $F^{n\times n}$ is a field if and only if $n = 1$.

[†]The name for the *Shoe-Sock Principle* comes from the identical property of inverse matrices, that is, $(AB)^{-1} = B^{-1}A^{-1}$. See Theorem 3.3.7 (ii) for further details.

"My happiness grows in direct proportion to my acceptance, and in inverse proportion to my expectations."
– Michael J. Fox

3.3 Matrix Inverses

When can we divide by a matrix? Does it even make sense to talk about matrix division? It depends on the matrix.

Definition 3.3.1. An $n \times n$ matrix A is **nonsingular** (or **invertible**) if there exists an $n \times n$ matrix B such that

$$AB = BA = I_n.$$

In this case, B is called an **inverse** of A. If A is not nonsingular, then it is **singular**.

Example 3.3.2. Let $A = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix}$ and $C = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix}$. Then

$$AC = \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and

$$CA = \begin{bmatrix} 5 & -3 \\ -3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Thus, A is invertible with inverse C. ∎

We should mention that inverses are unique. Suppose that A is invertible with inverses B and C. Thus,

$$B = BI_n = B(AC) = (BA)C = I_nC = C.$$

We will denote the inverse of A as A^{-1}. Therefore,

$$AA^{-1} = A^{-1}A = I_n.$$

Theorem 3.3.3. *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If* $\det(A) := ad - bc^{\dagger} \neq 0$, *then* A *is nonsingular with inverse*

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

If $ad - bc = 0$, *then* A *is singular.*

Example 3.3.4. Find the inverse of $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, if it exists.

Since $\det(A) = 1(4) - (2)3 = 4 - 6 = -2$, A is nonsingular, that is, it has an inverse, which is

$$A^{-1} = -\frac{1}{2} \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix}.$$ ∎

Example 3.3.5. Find the inverse of $A = \begin{bmatrix} 9 & 3 \\ 6 & 2 \end{bmatrix}$, if it exists.

Since $\det(A) = 9(2) - 3(6) = 18 - 18 = 0$, A has no inverse, that is, A is a singular matrix. ∎

For a nonsingular matrix A, the equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for all $\boldsymbol{b} \in F^n$ of the form $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.

Example 3.3.6. Solve the system of equations

$$\begin{cases} x_1 + 2x_2 &= 5 \\ 3x_1 + 4x_2 &= 6. \end{cases}$$

We seek to solve the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$ and A is the matrix from the previous equation. Thus,

$$A^{-1}(A\boldsymbol{x}) = A^{-1}\boldsymbol{b} \Rightarrow \boldsymbol{x} = A^{-1}\boldsymbol{b} = \begin{bmatrix} -2 & 1 \\ 3/2 & -1/2 \end{bmatrix} \begin{bmatrix} 5 \\ 6 \end{bmatrix} = \begin{bmatrix} -4 \\ 9/2 \end{bmatrix}.$$ ∎

Theorem 3.3.7. *Let A and B be $n \times n$ invertible matrices, let m be a positive integer, and r is a nonzero real number. Then A^{-1}, AB, $A^\top$, A^m, and rA are also invertible with:*

(i) $(A^{-1})^{-1} = A$ *(ii)* $(AB)^{-1} = B^{-1}A^{-1}$ *(iii)* $(A^\top)^{-1} = (A^{-1})^\top$

(iv) $(A^m)^{-1} = (A^{-1})^m := A^{-m}$ *(v)* $(kA)^{-1} = \frac{1}{k}A^{-1}$.

Note that (ii) is another example of the *Shoe-Sock Principle.*[††]

Corollary 3.3.8. *If A is invertible, then $AA^\top$ and $A^\top A$ are likewise invertible.*

Example 3.3.9. Inverse matrices allow us to solve matrix equations when otherwise we may have used division. For example, solve the equation $(AX^{-1})^{-1} + B = C$, assuming all matrices are $n \times n$ and nonsingular (when necessary).

We begin by adding $-B$ to both sides of the equation. This gives $(AX^{-1})^{-1} = C - B$. Using the Shoe-Sock Principle, we see that the left-hand side of the equation becomes $(X^{-1})^{-1}A^{-1} = XA^{-1} = C - B$. Finally, if we multiply the equation on both sides by A on the RIGHT side, we get

$$\begin{aligned} (XA^{-1})A &= (C - B)A \\ X(A^{-1}A) &= CA - BA \\ XI_n &= CA - BA \\ X &= CA - BA \end{aligned}$$

The side on which we are multiplying matters. Had we multiplied A on the left, we would have got AXA^{-1} which is not necessarily X. Likewise, $(XA^{-1})A \neq A(C - B)$ since multiplication by A must be done BOTH on the LEFT or BOTH on the RIGHT to guarantee equality (assuming all the products are even possible here). After all, multiplying by A on the LEFT is a different linear transformation than multiplying by A on the RIGHT. Do NOT assume matrices commute. Therefore, $X = \boxed{CA - BA}$.

Alternatively, we could solve for X in $(AX^{-1})^{-1} = C - B$ instead by taking inverses of both sides, that is, $[(AX^{-1})^{-1}]^{-1} = [C-B]^{-1}$, or $AX^{-1} = (C-B)^{-1}$. On the right-hand side, do not be tempted to distribute the exponent. While multiplication (and division) distribute over addition (and subtraction) even in the matrix algebraic setting, exponents do not. By analog, $(5-3)^{-1} = \dfrac{1}{5-3} = \dfrac{1}{2} \neq 5^{-1} - 3^{-1} = \dfrac{1}{5} - \dfrac{1}{3} = -\dfrac{2}{15}$. Continuing on, we will multiply both sides of the equation by A^{-1} on the LEFT. This gives $A^{-1}(AX^{-1}) = X^{-1} = A^{-1}(C-B)^{-1}$. Again taking the matrix inverse of both sides, we have $[X^{-1}]^{-1} = [A^{-1}(C-B)^{-1}]^{-1}$ or $X = (C-B)A = CA - BA$, agreeing with our solution from before. Of course, we used the Shoe-Sock Principle here at the end. This demonstrates that there could be more than one correct path toward the correct solution. Just make sure to adhere to all algebraic properties and restrictions, such as matrix multiplication being noncommutative and the Shoe-Sock Principle. ∎

Example 3.3.10. Solve the equation $B(XA)^{-1} = C$, assuming all matrices are $n \times n$ and nonsingular (when necessary).

We will demonstrate two paths to the solution of this matrix equation. Our first attempt begins with multiplying both sides of the equation by B^{-1} on the LEFT. This gives $B^{-1}(B(XA)^{-1}) = B^{-1}C \neq CB^{-1}$, or $(XA)^{-1} = B^{-1}C$. Inverting both sides then gives $XA = (B^{-1}C)^{-1} = C^{-1}[B^{-1}]^{-1} = C^{-1}B$. Lastly, we multiply both sides on the RIGHT by A^{-1}, giving $X = \boxed{C^{-1}BA^{-1}}$.

Alternatively, we could solve $B(XA)^{-1} = C$ by first multiplying by XA on the RIGHT of both sides. This gives $B(XA)^{-1}(XA) = C(XA)$, or $B = CXA$. Our goal is still to solve for X, however, notice that we cannot touch X in the current equation as it is sandwiched between two other matrices. In order to "unwrap," we must multiply by C^{-1} on the LEFT and then by A^{-1} on the RIGHT. This gives $XA = C^{-1}B$ and $XC^{-1}BA^{-1}$, respectively, agreeing with the solution of our first attempt. ∎

Theorem 3.3.11 (The Nonsingular Matrix Theorem)**.** *Let A be a square $n \times n$ matrix. Then the following statements are equivalent:*

(i) A *is a nonsingular matrix.*

(ii) A *is an invertible matrix, that is,* A *has a matrix inverse* A^{-1}*.*

(iii) There is an $n \times n$ *matrix* C *such that* $CA = I_n$*.*

(iv) There is an $n \times n$ *matrix* C *such that* $AC = I_n$*.*

(v) The equation $A\boldsymbol{x} = \boldsymbol{0}$ *has only the trivial solution.*

(vi) The linear system $A\boldsymbol{x} = \boldsymbol{b}$ *has no free variables for each* $\boldsymbol{b} \in F^n$*.*

(vii) The equation $A\boldsymbol{x} = \boldsymbol{b}$ *is consistent for each* $\boldsymbol{b} \in F^n$*.*

(viii) $A^\top$ *is an invertible matrix.*

(ix) A *is row equivalent to* I_n*.*

(x) A *has* n *pivot position.*

(xi) The rank of A *is* n*.*

(xii) The columns of A *span* F^n*.*

(xiii) The columns of A *form a linearly independent set.*

(xiv) The columns of A *form a basis for* F^n*.*

(xv) The rows of A *span* F^n*.*

(xvi) The rows of A *form a linearly independent set.*

(xvii) The rows of A *form a basis for* F^n*.*

(xviii) The linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ *is injective (one-to-one).*

(xix) The linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ *is surjective (onto).*

(xx) The linear transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ *is bijective (one-to-one and onto).*

Corollary 3.3.12. *Let A and B be $n \times n$ matrices. If AB is invertible, then A and B are likewise invertible.*

Example 3.3.13. Determine whether the following matrices are "nonsingular," "singular," or not enough information.

(a) Suppose that A is a 3×3 matrix over $\mathbb{C}$ such that $\text{nullity}(A) = 0$.

By part (vi) of the Nonsingular Matrix Theorem, we see that A is nonsingular since the nullity is equal to the number of free variables in the linear system $A\boldsymbol{x} = \boldsymbol{b}$.

(b) Suppose A is a 5×5 matrix over $\mathbb{Z}_2$. Furthermore, suppose that the equation $A\boldsymbol{x} = \mathbf{0}$ has 8 solutions.

By part (v) of the Nonsingular Matrix Theorem, we see that A is singular since the homogeneous equation $A\boldsymbol{x} = \mathbf{0}$ has nontrivial solutions.

(c) Suppose A is a 2×2 matrix such that

$$A \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

We do not have enough information to decide if A is nonsingular or not. On the one hand, the identity matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ satisfies this condition and is nonsingular. On the other hand, $\begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix}$ also satisfies this condition but is singular. ∎

Exercises

(Go to Solutions)

For Exercises 1-9, compute the inverse matrix of the given matrix using Theorem 3.3.3. Verify your inverse.

1. $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

3. $\begin{bmatrix} 9 & -4 \\ -7 & 5 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 2 & 5 \\ 3 & 7 \end{bmatrix}$

♠ 6. $\begin{bmatrix} 4 & 3 \\ -2 & 1 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 3 & 2 \\ 4 & 2 \end{bmatrix} \pmod 5$

♠ 8. $\begin{bmatrix} 10 & 2 \\ 5 & 3 \end{bmatrix} \pmod{11}$

♠ 9. $\begin{bmatrix} 1 & 1+i \\ -i & 2 \end{bmatrix}$

For Exercises 10-13, solve linear system $A\boldsymbol{x} = \boldsymbol{b}$ using the matrix inverse.

10. $A = \begin{bmatrix} 4 & 8 \\ 2 & 6 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

♠ 11. $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 1 \\ 2 & 1 & -5 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 0 \\ 3 \\ -2 \end{bmatrix}, A^{-1} = \begin{bmatrix} 4 & 13 & 5 \\ -3 & -11 & -4 \\ 1 & 3 & 1 \end{bmatrix}$

♠ 12. $A = \begin{bmatrix} 4 & 26 & 31 & -15 \\ -3 & -21 & -25 & 12 \\ 1 & 7 & 8 & -4 \\ 0 & -1 & -1 & 1 \end{bmatrix}, \boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, A^{-1} = \begin{bmatrix} 3 & 4 & 1 & 1 \\ -1 & 0 & 4 & 1 \\ 0 & -1 & -3 & 0 \\ -1 & -1 & 1 & 2 \end{bmatrix}$

♠ 13. $A \equiv \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{bmatrix}, \boldsymbol{b} \equiv \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, A^{-1} \equiv \begin{bmatrix} 2 & 1 & 0 \\ 2 & 1 & 2 \\ 2 & 2 & 0 \end{bmatrix} \pmod 3$

For Exercises 14-22, determine whether the following matrices are "nonsingular," "singular," or not enough information. Explain your reasoning.

♠ 14. A is a 4×4 matrix with $\text{rank}(A) = 4$.

♠ 15. A is a 5×4 matrix with $\text{rank}(A) = 4$.

♠ 16. A is a 3×3 matrix such that every vector in F^3 can be expressed as $A\boldsymbol{x}$, for some vector $\boldsymbol{x} \in F^3$.

♠ 17. A is a 5×5 matrix with a two rows of zeros in it row-reduced echelon form.

♠ 18. A is a 3×3 matrix in row-reduced echelon form.

♠ 19. A is a 3×3 matrix in row-reduced echelon form and a pivot in each column.

♠ 20. A is a 2×2 matrix such that $A\boldsymbol{x} = A\boldsymbol{y}$ for two distinct vectors $\boldsymbol{x}, \boldsymbol{y} \in F^2$.

♠ 21. A is row equivalent to a singular matrix.

♠ 22. A is a 3×3 matrix with columns vectors $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ such that $\boldsymbol{a}_1 - 2\boldsymbol{a}_2 = \boldsymbol{a}_3$.

For Exercises 23-28, solve matrix equations for the matrix X using matrix properties. You may assume all matrices are $n \times n$ and nonsingular. Be cautious! The order of multiplication matters!

23. $(BAX)^{-1} = C$

♠ 24. $AX + B = 0$

♠ 25. $A^{-1}(AX + C) = BA$

♠ 26. $(XA)^{-1}B = C$

♠ 27. $B(A + X)^{-1} = C$

28. $(DX - B)(CA)^{-1} = E$

29. Prove Corollary 3.3.8.

30. Prove Theorem 3.3.12.

[†]The value $\det(A) = ad - bc$ is called the **determinant** of A. We will learn more about determinants in Chapter 5.
[††]To remember this formula, think of the following proverb:

Put Your Socks On, Then Your Shoes
Take Your Shoes Off, Then Your Socks

"The opposite of love is not hate, it's indifference." – Elie Wiesel

3.4 Elementary Matrices

Definition 3.4.1. An **elementary matrix** is a matrix obtained by performing a single row operation on the identity matrix.

Since there are three types of row operations, there are three types of elementary matrices. Each elementary matrix is nonsingular, whose inverse is the elementary matrix corresponding to the inverse row operation.

1. (Replacement) If you replace Row i with Row i + cRow j in the identity matrix, then the resulting elementary matrix will have 1's across the main diagonal and 0's everywhere else except in the position (i, j) which will be c.

   For example, the matrix E_1, depicted below, corresponds to the row operation "replace Row 3 with Row 3 - 2Row 1."

   $$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

   Likewise, the matrix E_1^{-1} is the matrix which corresponds to the inverse row operation "replace Row 3 with Row 3 + 2Row 1."

   $$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$

2. (Interchange) If you interchange Row i with Row j in the identity matrix, then the resulting elementary matrix will be like the identity matrix except the ith and jth rows are interchanged.

   For example, the matrix E_2 corresponds to the row operation "interchange Rows 2 and 3."

   $$E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

   The inverse of E_2 is itself! It is also true that E_2 is equal to its own transpose.

   $$(E_2)^{-1} = E_2 = (E_2)^\top.$$

3. (Scaling) If you scale Row i by c in the identity matrix, then the resulting elementary matrix has 0's in all off-diagonal entries and 1's in the main diagonal entries except (i, i) which is c.

   For example, the matrix E_3 corresponds to the row operation "scale Row 2 by 7."

   $$E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   The inverse of E_3 corresponds to the row operation "scale Row 3 by 1/5."

   $$(E_3)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{7} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Example 3.4.2. Let $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Using the elementary matrices E_1, E_2, and E_3 from above, compute E_1A, E_2A, and E_3A.

$$E_1A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g-2a & h-2b & i-2c \end{bmatrix},$$

$$E_2A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}, \quad E_3A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 7d & 7e & 7f \\ g & h & i \end{bmatrix}$$

■

The previous example motivates the following proposition.

Proposition 3.4.3. *If an elementary row operation[†] is performed on an $m \times n$ matrix A, the resulting matrix can be written as EA, where the $m \times m$ matrix E is the elementary matrix associated to that elementary row operation.*

In the Nonsingular Matrix Theorem, we saw that an $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n. It turns out that the process of row-reducing A into I_n can produce the inverse matrix A^{-1} too.

Theorem 3.4.4 (Inversion Algorithm)**.** *Let A be an invertible $n \times n$ matrix A. Then any sequence of elementary row operations that reduces A to I_n also transforms I_n into A^{-1}.*

Proof. Suppose that A is invertible. Hence, $A \sim I_n$ by the Nonsingular Matrix Theorem. Then there exists some sequence of row operations transforming A into I_n. For each row operation in this sequence, there is a corresponding elementary matrix E_i. Say it took p row operations. Thus,

$$A \sim E_1A \sim E_2(E_1A) \sim \ldots \sim E_p(E_{p-1} \ldots E_1A) = I_n.$$

Thus, $(E_p \ldots E_2E_1)A = I_n$. Therefore, $E_p \ldots E_2E_1 = A^{-1}$. Since $(E_p \ldots E_2E_1)I_n = A^{-1}$, the sequence of row operations transforming A into I_n transforms I_n into A^{-1}. □

The Inversion Algorithm is a process for compute matrix inverses. Consider the matrix $[A \mid I_n]$. By row reduction,

$$[A \mid I_n] \sim [I_n \mid A^{-1}]. \tag{3.4.1}$$

Thus, row reduction saves the day again! Additionally, if $\mathcal{A}$ is the set of column vectors of A, which necessarily forms a basis for F^n, and $\mathcal{E}$ is the standard basis for F^n, then we see also that $A^{-1} = \underset{\mathcal{A} \leftarrow \mathcal{E}}{\mathcal{P}}$, the change of basis matrix from standard coordinates to $\mathcal{A}$-coordinates. Hence, the Inversion Algorithm is a special case of the Change-of-Basis Algorithm we saw in (2.8.1).

Example 3.4.5. Find the inverse of $A = \begin{bmatrix} 0 & 1 & -3 \\ 1 & -2 & 5 \\ -5 & 4 & 3 \end{bmatrix}$.

Using the method suggested above, we will row reduce A and apply these same row operations to I_n.

$$\left[\begin{array}{ccc|ccc} \boxed{0} & 1 & -3 & 1 & 0 & 0 \\ 1 & -2 & 5 & 0 & 1 & 0 \\ -5 & 4 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} \boxed{1} & -2 & 5 & 0 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 0 \\ -5 & 4 & 3 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -2 & 5 & 0 & 1 & 0 \\ 0 & \boxed{1} & -3 & 1 & 0 & 0 \\ 0 & -6 & 28 & 0 & 5 & 1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & -2 & 5 & 0 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 0 & \boxed{10} & 6 & 5 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -2 & 5 & 0 & 1 & 0 \\ 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & \frac{3}{5} & \frac{1}{2} & \frac{1}{10} \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -3 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & 1 & -3 & 1 & 0 & 0 \\ 0 & 0 & \boxed{1} & \frac{3}{5} & \frac{1}{2} & \frac{1}{10} \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & -2 & 0 & -3 & -\frac{3}{2} & -\frac{1}{2} \\ 0 & \boxed{1} & 0 & \frac{14}{5} & \frac{3}{2} & \frac{3}{10} \\ 0 & 0 & 1 & \frac{3}{5} & \frac{1}{2} & \frac{1}{10} \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{13}{5} & \frac{3}{2} & \frac{1}{10} \\ 0 & 1 & 0 & \frac{14}{5} & \frac{3}{2} & \frac{3}{10} \\ 0 & 0 & 1 & \frac{3}{5} & \frac{1}{2} & \frac{1}{10} \end{array}\right]$$

Therefore, $A^{-1} = \begin{bmatrix} \frac{13}{5} & \frac{3}{2} & \frac{1}{10} \\ \frac{14}{5} & \frac{3}{2} & \frac{3}{10} \\ \frac{3}{5} & \frac{1}{2} & \frac{1}{10} \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 26 & 15 & 1 \\ 28 & 15 & 3 \\ 6 & 5 & 1 \end{bmatrix}$. ■

A modification of the inversion algorithm can be used to factor a nonsingular matrix as product of elementary matrices. Suppose that A reduces to I_n via a sequence of p-many replacement row operations. Let $E_1, E_2, \ldots, E_p$ be the corresponding elementary matrices. Thus,

$$(E_p \ldots E_2 E_1)A = I_n, \quad \text{and} \quad A = (E_p \ldots E_2 E_1)^{-1} I_n = (E_1^{-1} E_2^{-1} \ldots E_p^{-1}).$$

Example 3.4.6. In Example 3.4.5 we reduced A into I_3 by the following sequence of row operations, given as descriptions and elementary matrices:

interchange Rows 1 and 2, $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

replace Row 3 with Row 3 + 5 Row 1, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 5 & 0 & 1 \end{bmatrix}$

replace Row 3 with Row 3 + 6 Row 2, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 6 & 1 \end{bmatrix}$

scale Row 3 by $\frac{1}{10}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{10} \end{bmatrix}$

replace Row 1 with Row 1 − 5 Row 3, $\begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

replace Row 2 with Row 2 + 3 Row 3, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$

replace Row 1 with Row 1 + 2 Row 2. $\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Following the same order but taking inverses, we get a factorization of A, namely:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -6 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{bmatrix} \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-6, for the row operation performed on an $m \times n$ matrix, write the corresponding elementary $m \times m$ matrix.

♠ 1. interchange Rows 1 and 3 in a 3×3 matrix.

♠ 2. scale Row 3 by 2 in a 4×4 matrix.

♠ 3. scale Row 2 by -5 in a 3×2 matrix.

♠ 4. replace Row 3 with Row $3 + 2$Row 1 in a 3×3 matrix.

♠ 5. replace Row 4 with Row $4 - 3$Row 2 in a 5×3 matrix.

♠ 6. replace Row 1 with Row $1 -$ Row 3 in a 4×2 matrix.

For Exercises 7-11, find the inverse matrix of the provided matrix, if possible. Verify your answer.

7. $\begin{bmatrix} 6 & 2 \\ 12 & 4 \end{bmatrix}$

8. $\begin{bmatrix} 4 & -3 & -1 \\ 1 & -1 & 0 \\ -2 & -1 & 2 \end{bmatrix}$

♠ 9. $\begin{bmatrix} 10 & -1 & -6 \\ -11 & 2 & 9 \\ -3 & 1 & 3 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 2 & 4 & 3 \\ 1 & 0 & 3 \\ 2 & 1 & 0 \end{bmatrix} \pmod 5$

11. $\begin{bmatrix} \frac{1}{2} & -1 & -\frac{3}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{3}{2} \\ -\frac{3}{2} & 3 & -\frac{7}{2} \end{bmatrix}$

For Exercises 12-15, factor the matrix as a product of elementary matrices. Answers may vary.

12. $\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$

♠ 13. $\begin{bmatrix} -6 & 8 & 13 \\ -3 & 4 & 6 \\ 1 & -1 & -2 \end{bmatrix}$

♠ 14. $\begin{bmatrix} 5 & -66 & -8 \\ -1 & 16 & 2 \\ 2 & -20 & -2 \end{bmatrix}$

15. $\begin{bmatrix} 3 & 1+3i & 7i \\ 2 & 1+2i & 4i \\ -2i & 2 & 7 \end{bmatrix}$

16. If

$$A \equiv \begin{bmatrix} 5 & 3 & 1 \\ 0 & 2 & 5 \\ 0 & 1 & 3 \end{bmatrix}, \; B \equiv \begin{bmatrix} 5 & 4 & 1 \\ 6 & 3 & 2 \\ 6 & 2 & 1 \end{bmatrix}, \; C \equiv \begin{bmatrix} 1 & 3 & 4 \\ 2 & 3 & 6 \\ 1 & 3 & 1 \end{bmatrix} \pmod 7,$$

solve the matrix equation $(AX)^{-1} \equiv B + C \pmod 7$.

[†]Elementary column operations can also be performed on A using the same three operations and the same three forms of elementary matrices, but the elementary matrices when multiplied on the right instead of left perform column operations.

"Whenever you're in conflict with someone, there is one factor that can make the difference between damaging your relationship and deepening it. That factor is attitude." – William James

3.5 Matrix Factorizations

In algebra, it is often important to be able to undo the process of multiplication, which is known as factorization. We see it in arithmetic, such as $6 = 2 \cdot 3$, and we see it in polynomials, $x^2 - 1 = (x-1)(x+1)$. Factorization can be extremely useful in solving problems where such objects arise, for example, when trying to solve the polynomial equation

$$x^2 - 1 = 0,$$

the subsequent factorization

$$(x-1)(x+1) = 0$$

illuminates the solution set $\{1, -1\}$. Such factorizations for matrices can be equally useful. Although no equivalent of prime numbers or irreducible polynomials exist for matrices, there are many very useful factorizations for matrices, much like the elementary factorizations we saw in the previous section. In this section, we discuss generalizations of the elementary matrices from the previous section and discussion their presence in matrix factorizations, in particular the LU factorization.

3.5.1 Generalizations of Elementary Matrices

Definition 3.5.1. A **diagonal matrix** is a square $n \times n$ matrix whose nondiagonal entries are all zero.

$$D = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}, \quad D^{-1} = \begin{bmatrix} \frac{1}{d_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{d_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{d_n} \end{bmatrix}.$$

The identity matrix I_n is a diagonal matrix all whose diagonal entries are 1. The zero matrix is a diagonal matrix all whose diagonal entries are 0. In a diagonal matrix, the diagonal entries need not be the same. In general, an $n \times n$ diagonal matrix D is of the form displayed to the right. Every diagonal matrix can be factored as a product of elementary matrices of scaling type (if we allow the possibility of zero scaling). As scaling type elementary matrices can have at most one non-unital diagonal entry, we can view diagonal matrices as their generalization. Similar to how scaling elementary matrices multiply, we see that if A is a matrix, then DA is the matrix where the ith row of A is scaled by d_i. Likewise, AD is the matrix where the jth column of A is scaled by d_j. In particular, a product of two diagonal matrices is a diagonal matrix whose diagonal entries are the products of the corresponding diagonal entries.

Furthermore, if no diagonal entry is zero, a diagonal matrix D is a product of elementary matrices and hence is invertible by the Nonsingular Matrix Theorem. The sum and scalar multiples of diagonal matrices are diagonal matrices. The set of diagonal matrices forms an n-dimensional subspace of $F^{n \times n}$. Additionally, the product of two diagonal matrices is diagonal.

Example 3.5.2. Let $A = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. Then A is a diagonal matrix and

$$A^{-1} = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{bmatrix}, \quad A^4 = \begin{bmatrix} \frac{1}{16} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 81 \end{bmatrix}, \quad A^{-4} = \begin{bmatrix} 16 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{81} \end{bmatrix}.$$

Furthermore, A can be factored into a product of scaling elementary matrices:

$$A = \begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}. \qquad \blacksquare$$

Definition 3.5.3. A diagonal matrix C is called a **scalar matrix** if all the diagonal entries are equal.

Essentially, scalar matrices all have the form $C = cI_n$ for some scalar c, which is why they get their name. Note that for any matrix A, we have that $CA = (cI_n)A = c(I_nA) = cA$, that is, multiplication by a scalar matrix is no different than scalar multiplication. Furthermore, scalar matrices are exactly those matrices which commute with all other matrices with regard to multiplication.

Definition 3.5.4. A **permutation matrix** is a square matrix formed by some rearrangement of the rows of I_n.

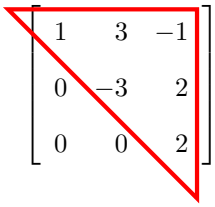
It can be shown that any permutation of objects can be accomplished by a sequence of transpositions, that is, a rearrangement of only two things at a time. This implies that a permutation matrix is a product of elementary matrices of interchange type. As such, a permutation matrix P is always non-singular. In fact, $P^{-1} = P^\top$. As an interchange elementary matrix has at most two rows permuted, we can view permutation matrices as their generalization. Similar to how interchange elementary matrices multiply, we see that if A is a matrix, then PA is the matrix where the rows of A are rearranged in the same way that the rows in I_n are rearranged for P. Likewise, AP is the matrix where the columns of A are rearranged in the same way that the columns in I_n are rearranged for P. In particular, a product of two permutation matrices is a permutation matrix, although the sum and scalar multiples of permutation matrices are no longer permutation matrices.

Example 3.5.5. Let $P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ is a permutation matrix. It can be easily verified that $P^\top = P^{-1}$. We also see that P factors as a product of interchange elementary matrices:

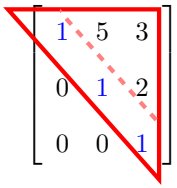
$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \blacksquare$$

Definition 3.5.6. An **upper triangular matrix** is a square $n \times n$ matrix whose entries below the main diagonal are all zero. A **lower triangular matrix** is a square $n \times n$ matrix whose entries above the main diagonal are all zero. A triangular that is either upper or lower triangular is called a **triangular matrix**. A matrix is **unit triangular** if it is triangular and all the diagonal entries are 1. A matrix is **strictly triangular** if it is triangular and all the diagonal entries are 0.

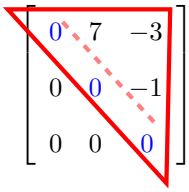
Upper Triangular Matrix



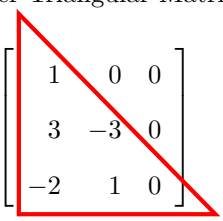
Unit Upper Triangular Matrix



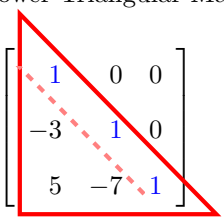
Strictly Upper Triangular Matrix



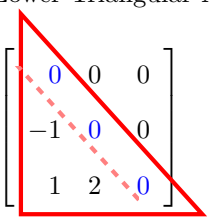
Lower Triangular Matrix



Unit Lower Triangular Matrix



Strictly Lower Triangular Matrix



In other words, an upper triangular matrix is a square matrix in echelon form, and a lower triangular matrix is just a square matrix in upside-down echelon form[†]. Of course, if a matrix is upper AND lower triangular, then it is actually a diagonal matrix.

An upper unit triangular matrix is a product of elementary matrices of replacement type, the kind used during the back-phase of Gauss–Jordan elimination and hence have non-zero entries above the diagonal. In general, an upper triangular matrix is a product of "backward" replacement elementary matrices and scaling (including by zero) elementary matrices. In particular, unit upper triangular matrices are those which are strictly a product of "backward" replacement elementary matrices, and hence can be viewed as their generalization. Upper triangular matrices are closed under sums and scalars, forming an $\frac{n(n-1)}{2}$-dimensional subspace of $F^{n\times n}$. Additionally, a product of upper triangular matrices is upper triangular, and an upper triangular matrix is invertible if and only if none of the diagonal entries are zero, in which case the inverse is likewise upper triangular. Furthermore, the transpose of an upper triangular matrix is lower triangular. Similar statements can be made about lower triangular matrices.

Example 3.5.7. Let $A = \begin{bmatrix} 2 & 0 & -5 \\ 0 & 1 & -3 \\ 0 & 0 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & -1 & -2 \\ 0 & 0 & -5 \\ 0 & 0 & 3 \end{bmatrix}$. Then A is nonsingular, but B is not by considering their diagonal entries. In particular,

$$A^{-1} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{5}{8} \\ 0 & 1 & \frac{3}{4} \\ 0 & 0 & \frac{1}{4} \end{bmatrix}, \quad AB = \begin{bmatrix} 6 & -2 & -19 \\ 0 & 0 & -14 \\ 0 & 0 & 12 \end{bmatrix}.$$

We also see that A factors as a product of replacement and scaling elementary matrices (in the case of B we allow for a zero-scaling matrix, which technically is not an elementary matrix):

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and}$$

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \blacksquare$$

3.5.2 The LU Factorization

Theorem 3.5.8. *Let A be an $m \times n$ matrix which has an echelon form which can be obtained solely by replacement row operations. Then there exists an $m \times n$ matrix U and an $m \times m$ matrix L such that*

$$A = LU,$$

U is an echelon form of A and L is a lower unit triangular matrix. In particular, the (i, j)-position of L for $i > j$ is $-c$ when the replacement $\text{Row}\, i \mapsto \text{Row}\, i + c\,\text{Row}\, j$ *was used during the row reduction process.*

Proof. When computing an echelon form of a matrix, the scaling row operation is never necessary. This one is only necessary for row-reduced echelon form. Thus, we need only use replacement or interchange to row reduce a matrix to echelon form. Suppose that A reduces to an echelon form U without interchange. Then there is a sequence of p-many replacement row operations transforming A into U. Let $E_1, E_2, \ldots, E_p$ be the corresponding elementary matrices. Thus,

$$(E_p \ldots E_2 E_1)A = U,$$

and

$$A = (E_p \ldots E_2 E_1)^{-1} U = (E_1^{-1} E_2^{-1} \ldots E_p^{-1})U.$$

Let $L = E_1^{-1} E_2^{-1} \ldots E_p^{-1}$. Now, each elementary matrix E_i is unit lower triangular. Its inverse E_i^{-1} is also unit lower triangular. The products of such matrices are also unit lower triangular, which gives the LU factorization.

□

The proof of this theorem provides us an algorithm for computing the LU factorization. **An Algorithm for finding an LU Factorization**

1. Reduce A to an echelon form U by a sequence of only forward-phase row replacement operations, if possible.

2. Place entries in L such that the same sequence of row operations reduces L to I_m, that is, place in the (i, j)-position the scalar $-c$ whenever $\text{Row}\, i \mapsto \text{Row}\, i + c\,\text{Row}\, j$ was used.

Example 3.5.9. Find an LU factorization of $A = \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 7 & 6 & 3 & 3 & 1 \\ 2 & 6 & 7 & 1 & 8 \\ 5 & 0 & 7 & 8 & 1 \end{bmatrix}$ (mod 11).

Since A has four rows, L will be 4×4. Below, you will see two columns of matrices. On the left, you will see the row reduction of A to an echelon form U using only forward-phase replacements. On the right, you will see the matrix L built brick-by-brick by placing the corresponding scalars below the diagonal.

$$A = \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 7 & 6 & 3 & 3 & 1 \\ 2 & 6 & 7 & 1 & 8 \\ 5 & 0 & 7 & 8 & 1 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \boxed{2} & 4 & 10 & 5 & 9 \\ 7 & 6 & 3 & 3 & 1 \\ 2 & 6 & 7 & 1 & 8 \\ 5 & 0 & 7 & 8 & 1 \end{bmatrix} \begin{matrix} \\ (\text{Row } 2 - 9\,\text{Row } 1) \\ (\text{Row } 3 - \text{Row } 1) \\ (\text{Row } 4 - 8\,\text{Row } 1) \end{matrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & * & 1 & 0 \\ 8 & * & * & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 0 & \boxed{3} & 1 & 2 & 8 \\ 0 & 2 & 8 & 7 & 10 \\ 0 & 1 & 4 & 1 & 6 \end{bmatrix} \begin{matrix} \\ \\ (\text{Row } 3 - 8\,\text{Row } 2) \\ (\text{Row } 4 - 4\,\text{Row } 2) \end{matrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 8 & 4 & * & 1 \end{bmatrix}$$

$$\sim \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 0 & 3 & 1 & 2 & 8 \\ 0 & 0 & 0 & \boxed{2} & 1 \\ 0 & 0 & 0 & 4 & 7 \end{bmatrix} \begin{matrix} \\ \\ \\ (\text{Row } 4 - 2\,\text{Row } 3) \end{matrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 8 & 4 & 2 & 1 \end{bmatrix}$$

$$\sim U = \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 0 & 3 & 1 & 2 & 8 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 8 & 4 & 2 & 1 \end{bmatrix}$$

Thus,

$$A = \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 7 & 6 & 3 & 3 & 1 \\ 2 & 6 & 7 & 1 & 8 \\ 5 & 0 & 7 & 8 & 1 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 8 & 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 0 & 3 & 1 & 2 & 8 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = LU \pmod{11}. \qquad \blacksquare$$

Consider the system of matrix equations depicted. Such a system might be considered, for example, if one needs to determine whether the list of vectors $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_p\}$ is in the span of the column vectors of A and, if so, what coefficients for a linear combination will satisfy the equations.

$$\begin{cases} A\boldsymbol{x} &= \boldsymbol{b}_1 \\ A\boldsymbol{x} &= \boldsymbol{b}_2 \\ & \vdots \\ A\boldsymbol{x} &= \boldsymbol{b}_p \end{cases}$$

Each equation $A\boldsymbol{x} = \boldsymbol{b}_i$ is solved using the SAME row operations to compute an echelon form of A. So computationally, it is often more efficient to record the necessary row operations once and for all. This is the case with past row-reductions such as $[\mathcal{B} \mid \mathcal{C}]$ or $[A \mid I_n]$. This leads to the LU factorization.

We have seen the convenience of solving an augmented matrix in echelon form (aka upper triangular matrix). Back substitution saves the day! Likewise, if a matrix is in lower triangular form, then it is essentially in "upside-down" echelon form and back substitution will lead to a quick solution of the augmented matrix. Therefore, suppose that

$$A = LU$$

is an LU factorization and consider the matrix equation

$$A\boldsymbol{x} = \boldsymbol{b}.$$

Let $\boldsymbol{y} = U\boldsymbol{x}$. Then

$$A\boldsymbol{x} = (LU)\boldsymbol{x} = L(U\boldsymbol{x}) = L\boldsymbol{y} = \boldsymbol{b}.$$

Thus, solving the equation $A\boldsymbol{x} = \boldsymbol{b}$ boils down to solving the two dramatically simpler matrix equations

$$L\boldsymbol{y} = \boldsymbol{b} \quad \text{and} \quad U\boldsymbol{x} = \boldsymbol{y}$$

Example 3.5.10. We saw in the previous example that

$$A = \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 7 & 6 & 3 & 3 & 1 \\ 2 & 6 & 7 & 1 & 8 \\ 5 & 0 & 7 & 8 & 1 \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 9 & 1 & 0 & 0 \\ 1 & 8 & 1 & 0 \\ 8 & 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 10 & 5 & 9 \\ 0 & 3 & 1 & 2 & 8 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} = LU \pmod{11}.$$

Using this LU factorization of A, solve $A\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{b} = (1, 2, 3, 4)$.

We first solve the matrix equation $L\boldsymbol{y} = \boldsymbol{b}$, which corresponds to the augmented matrix:

$$\left[\begin{array}{ccccc|c} \boxed{1} & 0 & 0 & 0 & 1 \\ 9 & 1 & 0 & 0 & 2 \\ 1 & 8 & 1 & 0 & 3 \\ 8 & 4 & 2 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 & 4 \\ 0 & 8 & 1 & 0 & 2 \\ 0 & 4 & 2 & 1 & 7 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & 2 & 1 & 2 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & 7 \end{array}\right].$$

Thus, $\boldsymbol{y} = (1, 4, 3, 7)$.

Next, we need to solve the equation $U\boldsymbol{x} = \boldsymbol{y}$:

$$\left[\begin{array}{ccccc|c} 2 & 4 & 10 & 5 & 9 & 1 \\ 0 & 3 & 1 & 2 & 8 & 4 \\ 0 & 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & \boxed{5} & 7 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 4 & 10 & 5 & 9 & 1 \\ 0 & 3 & 1 & 2 & 8 & 4 \\ 0 & 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 & \boxed{1} & 8 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 4 & 10 & 5 & 0 & 6 \\ 0 & 3 & 1 & 2 & 0 & 6 \\ 0 & 0 & 0 & \boxed{2} & 0 & 6 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 4 & 10 & 5 & 0 & 6 \\ 0 & 3 & 1 & 2 & 0 & 6 \\ 0 & 0 & 0 & \boxed{1} & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right]$$

$$\sim \left[\begin{array}{ccccc|c} 2 & 4 & 10 & 0 & 0 & 2 \\ 0 & \boxed{3} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 2 & 4 & 10 & 0 & 0 & 2 \\ 0 & \boxed{1} & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right] \sim \left[\begin{array}{ccccc|c} \boxed{2} & 0 & 5 & 0 & 0 & 2 \\ 0 & 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 8 & 0 & 0 & 1 \\ 0 & 1 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 8 \end{array}\right]$$

Thus, $\boldsymbol{x} = (1, 0, 0, 3, 8) + t(3, 7, 1, 0, 0)$ is the general solution.

We should mention that the above example required 27 arithmetic operations while the usual row-reduction method would have required 46 operations. ∎

To handle row interchanges, the LU factorization above can be modified easily to produce an PLU-factorization, where P is a permutation matrix. Similarly, to handle row scaling, a diagonal matrix D can be introduced to given instead the LDU-factorization. Combining all three row operations together leads to the $PLDU$-factorization. The corresponding system of equation of each of these factorizations is still relatively easy to solve.

Exercises

(Go to Solutions)

For Exercises 1-4, determine if the matrix is a scalar matrix.

1. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

2. $\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

QUICK! For Exercises 5-7, multiply by the diagonal matrices using the discussion following Definition 3.5.1 in LESS THAN 10 SECONDS!

♠ 5. $\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -3 & 2 \\ 1 & 5 \end{bmatrix}$

♠ 6. $\begin{bmatrix} -3 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 5 \\ 2 & -2 & -3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 3 & -1 & 4 \\ -2 & 1 & -5 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

For Exercises 8-8, find A^2, A^3, A^{-1}, and A^{-3}.

♠ 8. $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$.

For Exercises 9-9, the permutation matrix P will be listed first and another matrix A will be listed second. Factor P as a product of interchange elementary matrices (answers may vary). Compute PA and explain how the rows of A have been permuted. Compute AP and explain how the rows of A have been permuted.

9. $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 5 & 8 & 10 \\ 5 & 2 & 6 & 9 \\ 8 & 6 & 3 & 7 \\ 10 & 9 & 7 & 4 \end{bmatrix}$

For Exercises 10-11, factor the following matrices as a product of elementary matrices.

♠ 10. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 4 & 1 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 3 & 12 & -6 \\ 0 & 5 & 45 \\ 0 & 0 & 2 \end{bmatrix}$

For Exercises 12-13, find the LU-factorization of the provided matrix.

12. $\begin{bmatrix} 2 & 4 & -1 \\ 8 & 6 & 4 \\ 6 & 8 & 10 \end{bmatrix}$

13. $\begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

For Exercises 14-18, find the LU-factorization of the coefficient matrix A below and use this factorization to solve the linear system as in Example 3.5.10. [**Hint**: Remember to NOT scale NOR interchange.]

♠ 14. $\begin{bmatrix} 2 & -5 \\ -4 & 13 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$

♠ 15. $\begin{bmatrix} 2 & -2 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$

16. $\begin{bmatrix} 2 & 6 & 0 \\ 16 & 51 & 0 \\ 12 & 42 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \\ 15 \end{bmatrix}$

♠ 17. $\begin{bmatrix} 2 & 2 & -2 \\ -4 & -3 & 4 \\ -1 & -2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -4 \\ -2 \\ 6 \end{bmatrix}$

♠ 18. $\begin{bmatrix} 2 & 1 & 2 \\ 6 & 6 & 0 \\ 4 & 5 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \equiv \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \pmod 7$

19. Show that if A is a lower triangular matrix then $A^\top$ is upper triangular.

[†] I suppose it could be called a *chandelier form*.

"Life is a mirror and will reflect back to the thinker what he thinks into it." – Ernest Holmes

3.6 Linear Transformations on $\mathbb{R}^2$

In this section, we will explore how multiplication by a nonsingular matrix transforms geometry in $\mathbb{R}^2$. While analogous statements can be extended to the vector space F^n, based upon the field F itself, only some geometric interpretations may apply.

Let $a, b \in \mathbb{R}$ such that $a, b > 0$.

We can **stretch** $\mathbb{R}^2$ **horizontally** by multiplying the (scaling) matrix

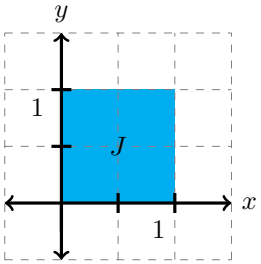
$$\begin{bmatrix} a & 0 \\ 0 & 1 \end{bmatrix}.$$

We can **stretch** $\mathbb{R}^2$ **vertically** by multiplying the matrix

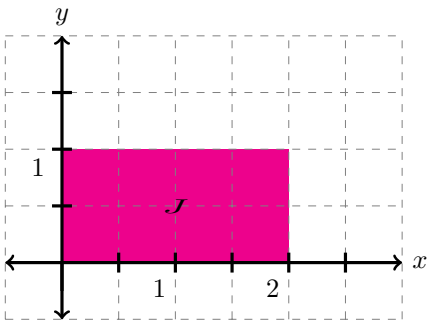
$$\begin{bmatrix} 1 & 0 \\ 0 & b \end{bmatrix}.$$

Thus, multiplication by a diagonal matrix will rescale the x- and y-axes by the diagonal entries. When $a < 1$ or $b < 1$, we say the geometry is **compressed** since the scale becomes smaller than the original.

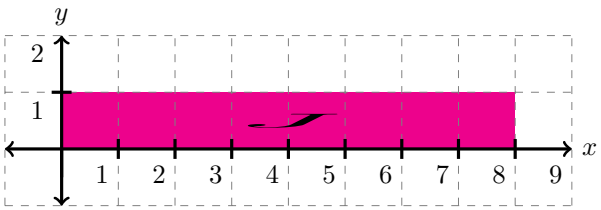
Example 3.6.1. The **unit square** J is the rectangle with vertices at $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$. The unit square is displayed below in cyan. It will be a useful tool to visualize how matrix multiplication distorts the geometry by seeing how a matrix transforms the unit square J.



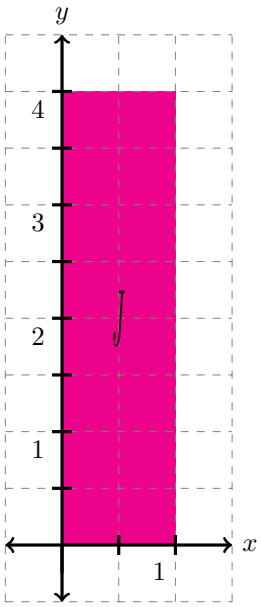
(a) Multiplication by the matrix $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$ horizontally stretches the plane by a factor of two.



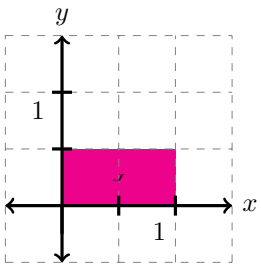
(b) Multiplication by the matrix $\begin{bmatrix} 8 & 0 \\ 0 & 1 \end{bmatrix}$ horizontally stretches the plane by a factor of eight.



(c) Multiplication by the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$ vertically stretches the plane by a factor of four.



(d) Multiplication by the matrix $\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$ vertically comprresses the plane by a factor of two.



■

Example 3.6.2. The matrix $\begin{bmatrix} 1/2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$ will vertically stretch the four vertices of J by a factor of 3 and will horizontally compress the vertices by a factor of 2 (or we can say that they are horizontally stretched by a factor of 1/2). In particular,

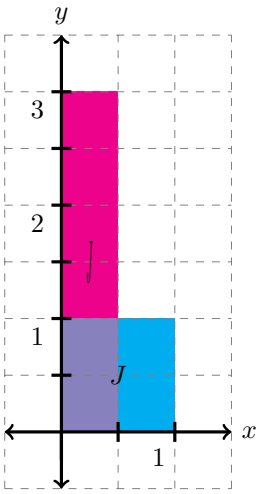
$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 3 \end{bmatrix}$$

The image of J is displayed to the right in magenta.



■

In plane geometry, a shear mapping (or transvection) is a linear map that displaces each point in fixed direction, by an amount proportional to its signed distance from a line that is parallel to that direction. Let $m \in \mathbb{R}$.

We can **shear** $\mathbb{R}^2$ **horizontally** by multiplying the (replacement) matrix

$$\begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}.$$

We can **shear** $\mathbb{R}^2$ **vertically** by multiplying the matrix

$$\begin{bmatrix} 1 & 0 \\ m & 1 \end{bmatrix}.$$

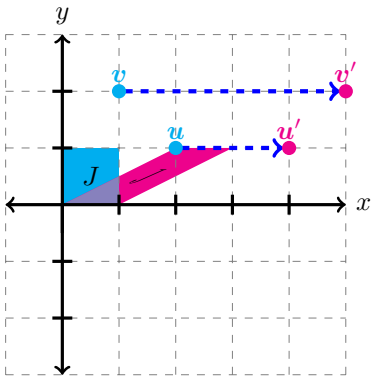
Example 3.6.3. Consider the shearing of the unit square and the points $\boldsymbol{u} = (2, 1)$ and $\boldsymbol{v} = (1, 2)$:

(a) horizontally by a factor of 2.

The images of these points under this shear are

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

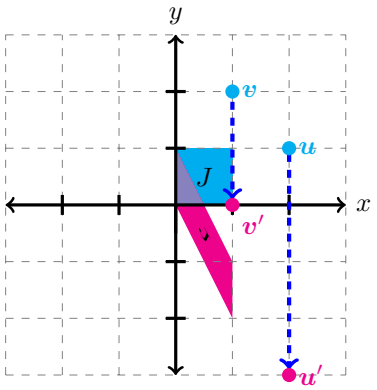


(b) vertically by a factor of -2.

The images of $\boldsymbol{u}$ and $\boldsymbol{v}$ under this shear are

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -3 \end{bmatrix};$$

$$\begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

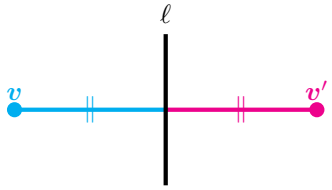


■

Reflections across the x- and y-axis are accomplished by multiplying by the scaling (diagonal) matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

respectively.



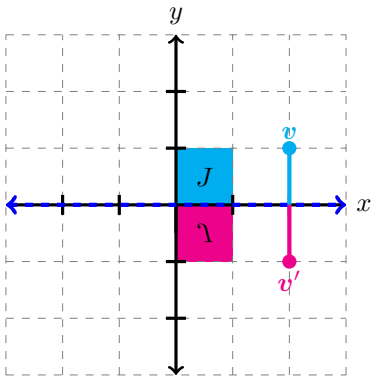
The interchange (permutation) matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ produces a reflection across the line $y = x$. The composite of reflections across the x- and y-axis gives a reflection through the origin. General reflections will be discussed in the homework.

Example 3.6.4. Consider the reflections of point $\boldsymbol{v} = (2, 1)$ across:

(a) the x-axis.

The image of $\boldsymbol{v}$ under this reflection is

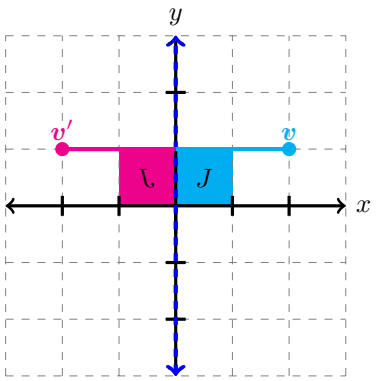
$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$



(b) the y-axis.

The image of $\boldsymbol{v}$ under this reflection is

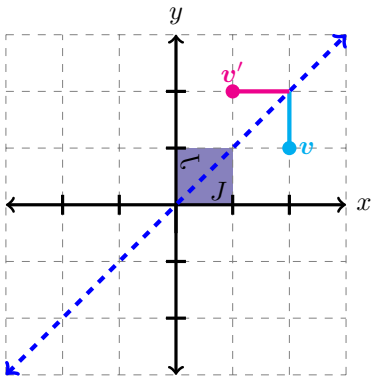
$$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$



(c) $y = x$.

The image of $\boldsymbol{v}$ under this reflection is

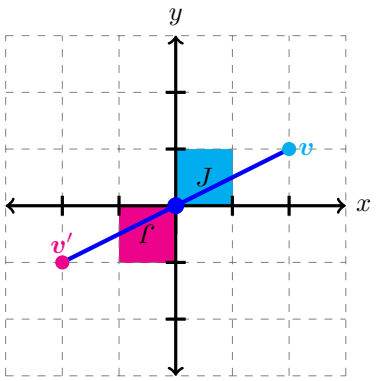
$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$



(d) the origin.

The image of $\boldsymbol{v}$ under this reflection is

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix}.$$



■

Theorem 3.6.5. *Let A be a 2×2 nonsingular matrix. Then the matrix transformation associated to A geometrically is a composition of shears, reflections, and stretches/compressions.*

The following example illustrates Theorem 3.6.5.

Example 3.6.6. Let $A = \begin{bmatrix} -4 & -5 \\ 2 & 4 \end{bmatrix}$. We reduce A to its row reduced echelon form in the following way:

$$\begin{bmatrix} -4 & -5 \\ 2 & 4 \end{bmatrix} \sim \begin{bmatrix} 2 & 4 \\ -4 & -5 \end{bmatrix} \begin{matrix} \\ \text{(Row 2 + 2 Row 1)} \end{matrix} \sim \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} \begin{matrix} \\ \left(\frac{1}{2}\text{Row 1}\right) \end{matrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{matrix} \\ \left(\frac{1}{3}\text{Row 2}\right) \end{matrix}$$

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{matrix} \\ \text{(Row 1 - 2 Row 2)} \end{matrix} \sim \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Following this sequence of row operation gives the following elementary factorization of A:

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

To translate this elementary factorization into a geometric interpretation, we read the factorization right-to-left. Why? Because with the transformation $\boldsymbol{x} \mapsto A\boldsymbol{x} = (E_1E_2E_3E_4E_5)\boldsymbol{x}$, assuming E_i denotes the ith elementary factor in the above factorization of A, the first matrix to transform $\boldsymbol{x}$ will be E_5. The next to transform the resultant vector would be E_4, and onward toward the left. Hence, multiplication by A has the affect of shearing horizontally by a factor of 2, stretching vertically by a factor of 3, stretching horizontally by a factor of 2, shearing vertically by a factor of -2, and reflecting across the diagonal line $y = x$. ∎

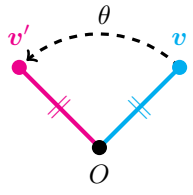
Example 3.6.7. Let $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$. We can row reduce A by interchanges Rows 1 and 2, scaling Row 1 by $\frac{1}{2}$, and replacing Row 1 with Row $1 - \frac{1}{2}$ Row 2. This gives the following factorization into elementary matrices as

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 0 & 1 \end{bmatrix}.$$

Therefore, A corresponds to the transformation shear horizontally by $\frac{1}{2}$, horizontally stretch by 2, and reflect across the line $y = x$. ∎

A counterclockwise rotation in $\mathbb{R}^2$ by angle θ is captured by the matrix

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$



For a clockwise rotation by θ, take the previous matrix's inverse, $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$.

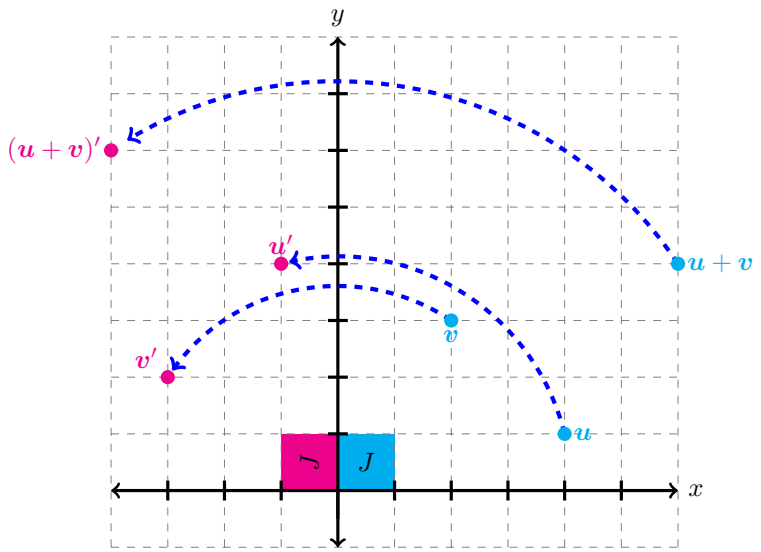
Example 3.6.8. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given as counterclockwise rotation by $\frac{\pi}{2}$ (or 90°). Find the images under T of $\boldsymbol{u} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, and $\boldsymbol{u} + \boldsymbol{v} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$.

Note that $[T] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. Then

$$T(\boldsymbol{u}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$$

$$T(\boldsymbol{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$$

$$T(\boldsymbol{u}+\boldsymbol{v}) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 6 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$



■

A similar analysis may be used to describe the geometry of matrix transformation in $\mathbb{R}^3$ (and higher).

Exercises

(Go to Solutions)

For Exercises 1-5, find the matrix for the operator that performs the stated succession of geometric transformations.

♠ 1. Compress by a factor of 2 in the x-direction, then expands by a factor of 5 in the y-direction.

♠ 2. Expands by a factor of 5 in the y-direction, then shears by a factor of 2 in the y-direction.

♠ 3. Reflects about $y = x$, then rotates through an angle of π radians about the origin.

♠ 4. Reflects about the y-axis, then expands by a factor of 5 in the x-direction, and then reflects about $y = x$.

♠ 5. Rotates through $\dfrac{\pi}{6}$ about the origin, then shears by a factor of -2 in the y-direction, and then expands by a factor of 3 in the y-direction.

For Exercises 6-9, express the matrix as a product of elementary matrices, and then describe the effect of multiplication by this matrix in terms of shears, compressions/stretches, and reflections. Draw the image of the unit square under this matrix transformation. Answers may vary.

♠ 6. $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$ ♠ 7. $\begin{bmatrix} 1 & 0 \\ 0 & -8 \end{bmatrix}$ ♠ 8. $\begin{bmatrix} 0 & -2 \\ 4 & 0 \end{bmatrix}$ ♠ 9. $\begin{bmatrix} 3 & -6 \\ 6 & 1 \end{bmatrix}$

10. What matrix, if multiplied by, would have the affect to the plane $\mathbb{R}^2$ of reflection across the line $y = -x$? How could you factor this matrix using the matrix transformation discussed in this section?

♠ 11. Let $A = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix}$, where $\theta > 0$.

(a) Show that $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1}$.

(b) For $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1}$ describe the sequence of geometric transformations that correspond to this matrix factorization. Conclude that the geometric transformation corresponding to A is reflection across the line ℓ through the origin which forms an angle θ with the x-axis.

(c) For $\theta = 0$, $\dfrac{\pi}{4}$, and $\dfrac{\pi}{2}$, show that this agrees with the three reflections discussed in this section.

(d) For $\theta = \dfrac{\pi}{6}$, compute the reflection matrix and draw the image of the unit square under this matrix transformation.

(e) Modifying part (b), find a 2×2 matrix B where multiplication by B performs the geometric transformation of shearing by a factor of m in the direction of the line ℓ through the origin which forms an angle θ with the x-axis.

12. Prove Theorem 3.6.5.

"The people who oppose your ideas are inevitably those who represent the established order that your ideas will upset." – Anthony J. D'Angelo

3.7 Representations of Linear Transformations as Matrices

We have seen before that linear transformations are those maps between vector spaces which preserve the vector structure, that is, preserve vector addition and scalar multiplication. We have seen that matrix multiplication is a linear transformation, hence all the geometric transformations in the previous section are linear transformations. We will now see that all linear transformations are essentially just matrix multiplication.

Example 3.7.1. Let $\boldsymbol{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation such that

$$T(\boldsymbol{e}_1) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{and} \quad T(\boldsymbol{e}_2) = \begin{bmatrix} 3 \\ -5 \\ 0 \end{bmatrix}.$$

With no additional information, find a formula for the image of an arbitrary vector $\boldsymbol{x} \in \mathbb{R}^2$.

Since $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2$ and since T is a linear transformation, it must hold that

$$\begin{aligned} T(\boldsymbol{x}) &= T(x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2) = x_1T(\boldsymbol{e}_1) + x_2T(\boldsymbol{e}_2) \\ &= x_1\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + x_2\begin{bmatrix} 3 \\ -5 \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 + 3x_2 \\ 2x_1 - 5x_2 \\ 3x_1 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & -5 \\ 3 & 0 \end{bmatrix}\boldsymbol{x}. \end{aligned}$$

In particular, T can be represented as a matrix transformation. ∎

The argument applied in the previous example works in general.

Theorem 3.7.2. *Let $\boldsymbol{e}_i$ be the vector in F^n such that the ith entry is 1 and all other entries are 0. Let $T : F^n \to F^m$ be a linear transformation and let $A = \begin{bmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \dots & T(\boldsymbol{e}_n) \end{bmatrix}$. Let $\boldsymbol{x} \in F^n$. Then*

$$T(\boldsymbol{x}) = A\boldsymbol{x}.$$

*This matrix A is called the **standard matrix for the linear transformation** T and is often denoted $[T]$.*

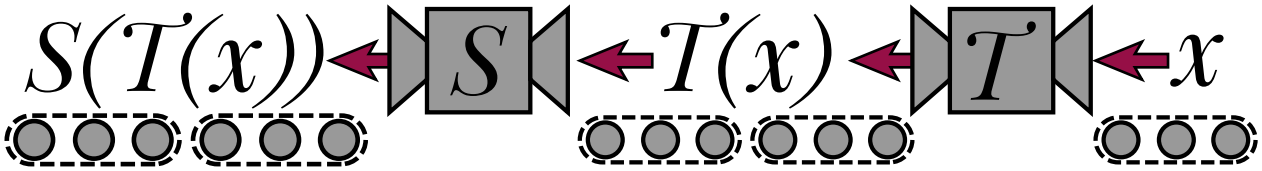
Example 3.7.3. Find the standard matrix A for the linear transformation $T(x_1,\ x_2) = (3x_1 + x_2,\ 2x_1 - 4x_2)$.

Since $T(\boldsymbol{e}_1) = (3,\ 2)$ and $T(\boldsymbol{e}_2) = (1,\ -4)$, we have that $[T] = A = \begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}$ and

$$T(\boldsymbol{x}) = A\boldsymbol{x} = \begin{bmatrix} 3 & 1 \\ 2 & -4 \end{bmatrix}\boldsymbol{x}.$$

∎

Recall that if $S : F^m \to F^p$ and $T : F^n \to F^m$ are functions, then their composite $S \circ T$ is the function given by the rule $(S \circ T)(\boldsymbol{x}) = S(T(\boldsymbol{x}))$ where $\boldsymbol{x} \in F^n$. Functions are often viewed, by parable, as machines on a conveyor belt. As an element $\boldsymbol{x}$ in the domain of T rolls into the machine T, it is transformed into the element $T(\boldsymbol{x})$. The composite $S \circ T$ is then the machine formed by placing the machines S and T next each other on the conveyor belt. The following theorem states that matrix representations are compatible with function composition.



Theorem 3.7.4. *Let $S : F^m \to F^p$ and $T : F^n \to F^m$ be linear transformations with standard matrices A and B, respectively. Let $\boldsymbol{x} \in F^n$. Then*

$$S \circ T(\boldsymbol{x}) = S(T(\boldsymbol{x})) = AB\boldsymbol{x},$$

that is, the matrix which represents the composite of two linear transformations is the product of their matrix representations.

Recall that a linear transformation $T : F^n \to F^m$ is *one-to-one* if and only if $T(\boldsymbol{x}) = T(\boldsymbol{y})$ implies that $\boldsymbol{x} = \boldsymbol{y}$ if and only if $\ker T = \{\mathbf{0}\}$, that is, the only vector mapping to zero is zero itself. Given the matrix representation $[T] = A$, we see that $\ker T = A$, that is, the vectors mapped to $\mathbf{0}$ by T are exactly those vectors whose product with A is $\mathbf{0}$. Hence, finding the null space of A is equivalent to finding the kernel of T. In particular, a linear transformation is one-to-one if and only if the columns of its matrix representation are linearly independent.

Example 3.7.5. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a transformation given by the rule

$$T(x, y) = (x + y, 0, 2x + 3y).$$

Then

$$[T] = \begin{bmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 2 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 \\ 2 & 3 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \\ 0 & 0 \end{bmatrix}.$$

The row-reduced echelon form of $[T]$ shows that the columns of $[T]$ are linear independent since there is a pivot in each column. Therefore, T is one-to-one, that is, $\ker T = \{\mathbf{0}\}$. ∎

Example 3.7.6. Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear transformation given by the rule

$$S(x, y, z) = (x + y - 2z, -y + z).$$

Then

$$[S] = \begin{bmatrix} S(\boldsymbol{e}_1) & S(\boldsymbol{e}_2) & S(\boldsymbol{e}_3) \end{bmatrix} = \begin{bmatrix} 1 & 1 & -2 \\ 0 & -1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & -1 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & -1 \end{bmatrix}.$$

The row-reduced echelon form of $[S]$ shows that the columns of $[S]$ are linear dependent since there is no pivot in the third column. Therefore, S is not one-to-one. In particular, $\ker S = \text{Span}\{[1, 1, 1]^\top\}$. ∎

The homogeneous system $A\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution when A has a non-pivot column. The nullity of A is the number of non-pivot columns in A. If A is an $m \times n$ matrix and $m < n$, that is, A has more columns than rows, then A necessarily has non-pivot columns. Hence, the nullity$(A) \geq 1$.

Recall that a linear transformation $T : F^n \to F^m$ is *onto* if and only if for all vectors $\boldsymbol{b} \in F^m$ there exists a vector $\boldsymbol{x} \in F^n$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$ if and only if $\text{im}(T) = F^m$, that is, the set of images of T is the whole target space F^m. Given the matrix representation $[T] = A$, we see that $\text{im}\, T = \text{Col}\, A$, that is, the vectors which are linear combinations of the columns of A are exactly the vectors which can come out of T. Hence, finding the column space of A is equivalent to finding the image of T. In particular, a linear transformation is onto if and only if the columns of its matrix representation span F^m.

Example 3.7.7. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the transformation given in Example 3.7.5. The row-reduced echelon form of $[T]$ shows that $[T]$ has no pivot in the third row. Therefore, T is not onto. In particular, $[0, 1, 0]^\top \notin \text{im}\, T$. ∎

Example 3.7.8. Let $S : \mathbb{R}^3 \to \mathbb{R}^2$ be the transformation given in Example 3.7.6. The row-reduced echelon form of $[S]$ shows that $[S]$ has a pivot in each row. Therefore, S is onto. In particular, $\text{im}\, S = \mathbb{R}^2$. ∎

The inconsistency of the equation $A\boldsymbol{x} = \boldsymbol{b}$ may occur when A has a row of zeros in echelon form but the corresponding position in the augmented column is nonzero. In this case, consistency is dependent on the choice of $\boldsymbol{b}$ and at least one choice of $\boldsymbol{b}$ will make $A\boldsymbol{x} = \boldsymbol{b}$ inconsistent. These rows of zeros in the echelon form correspond to non-pivot rows of A. Let the **conullity** of A denote the number of non-pivot rows of A. If A is an $m \times n$ matrix and $m > n$, that is, A has more rows than columns, then A necessarily has non-pivot rows. Hence, the conullity$(A) \geq 1$.

Proposition 3.7.9. *Let $T : F^n \to F^m$ is a linear transformation.*

(xxi) If $m > n$, that is, $[T]$ has more rows than columns, then T cannot be onto.

(xxii) If $m < n$, that is, $[T]$ has more columns than rows, then T cannot be one-to-one.

Exercises

(Go to Solutions)

For Exercises 1-5, for the linear transformations T find its standard matrix $[T]$.

♠ 1. $T : \mathbb{R}^2 \to \mathbb{R}^2 : T(x_1, x_2) = (2x_1 - x_2, -4x_2)$.

♠ 2. $T : \mathbb{R}^3 \to \mathbb{R}^4 : T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 + x_3 \\ 3x_1 + 14x_3 - 5x_2 \\ 3x_2 - 5x_1 - 18x_3 \\ 19x_2 - 7x_1 - 40x_3 \end{bmatrix}$.

♠ 3. $T : \mathbb{C} \to \mathbb{C}^4 : T(z) = (z, iz, -z, -iz)$.

♠ 4. $T : \mathbb{Z}_2^4 \to \mathbb{Z}_2 : T(x_1, x_2, x_3, x_4) = x_1 + x_2 + x_3 + x_4$.

♠ 5. $T : \mathbb{Z}_5^3 \to \mathbb{Z}_5^3 : T(x_1, x_2, x_3) = (x_3, x_2, x_1)$.

For Exercises 6-6, find the standard matrix of the composite function $S \circ T$.

♠ 6. $T : \mathbb{R}^2 \to \mathbb{R}^3 : T(x_1, x_2) = (2x_2, 3x_1, 2x_1 - 3x_2)$ and $S : \mathbb{R}^3 \to \mathbb{R}^2 : S(y_1, y_2, y_3) = (y_3 - 2y_2, y_1 - 2y_3)$

For Exercises 7-11, determine if the given linear transformation from Exercises 1-5 is one-to-one or onto. Find a basis for its kernel and for its image.

♠ 7. Exercise 1 ♠ 8. Exercise 2 ♠ 9. Exercise 3 ♠ 10. Exercise 4 ♠ 11. Exercise 5

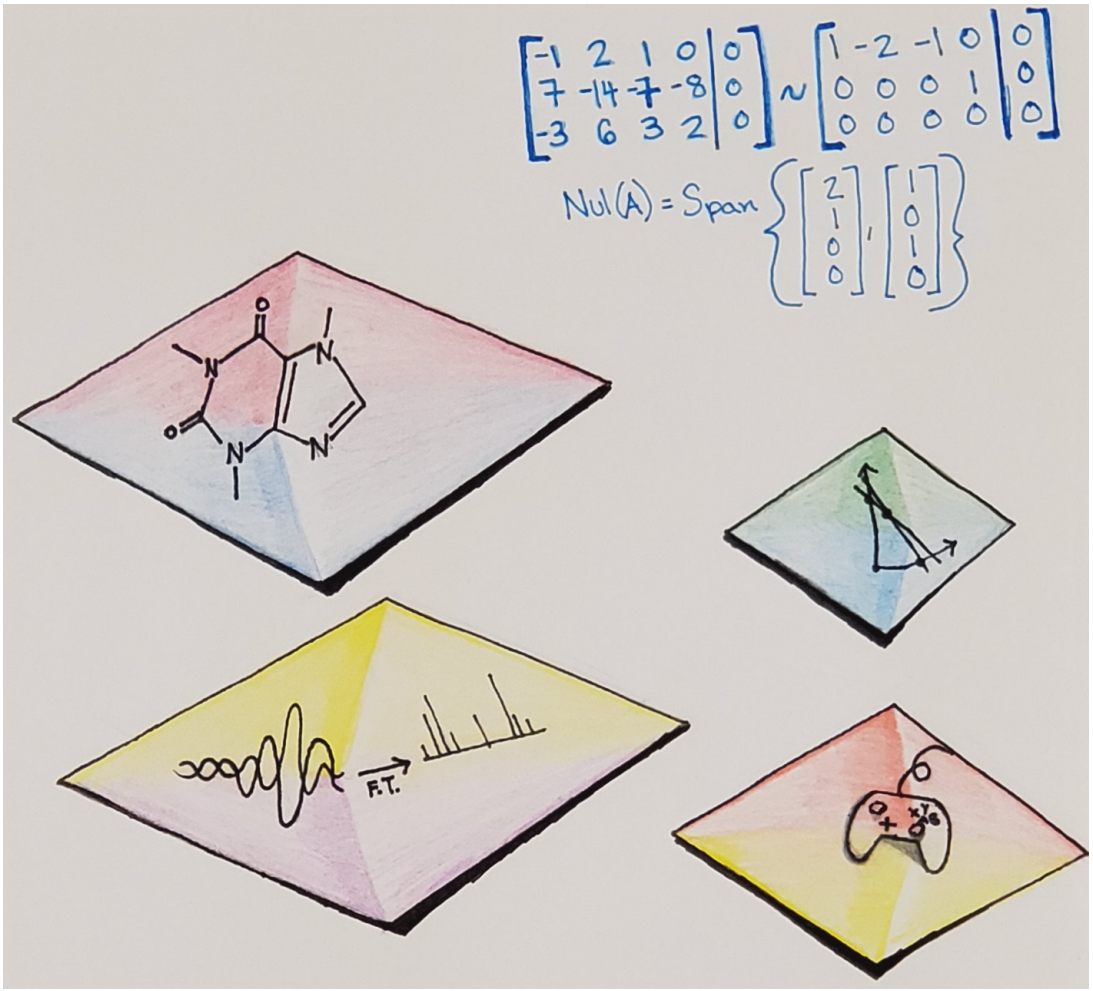
QUICK! For Exercises 7-11, determine in less than 10 seconds if the linear transformation CANNOT be one-to-one and if it CANNOT be onto.

12. $T : \mathbb{R}^3 \to \mathbb{R}^3 : T(x, y, z) = (x, y, z)$
13. $T : \mathbb{R}^3 \to \mathbb{R}^2 : T(x_1, x_2, x_3) = (x_1 + 2x_2, x_3 - 3x_2)$
14. $T : \mathbb{R}^3 \to \mathbb{R}^4 : T(x_1, x_2, x_3) = (2x_1 + 3x_3, x_2 + 4x_1 + x_3, x_3 + x_1 + x_2, x_1 + 3x_2 + x_3)$
15. $T : \mathbb{C} \to \mathbb{C}^3 : T(z) = (z, z, z)$
16. $T : \mathbb{C} \to \mathbb{C}^3 : T(z) = ((3 - i)z, (-2 + 2i)z, z)$
17. $T : \mathbb{Z}_5^3 \to \mathbb{Z}_5 : T(x, y, z) = x + y + z$
18. $T : \mathbb{Z}_7^2 \to \mathbb{Z}_7^2 : T(x, y) = (4x + 2y, 3y)$

Chapter 4

Orthogonality

In this chapter, we will begin to develop the geometric framework of lengths and angles through the lens of linear algebra. While the notions of affine geometry were introduced earlier in Chapter 2, this chapter will develop the notions of Euclidean geometry. As such, this geometric development will not be entirely possible for the finite fields $\mathbb{Z}_p$ we have often considered, although some notion of orthogonality are available for general fields. In this chapter the only fields F we will consider for scalars will be $\mathbb{R}$ and $\mathbb{C}$.



"Success isn't measured by money or power or social rank. Success is measured by your discipline and inner peace." – Mike Ditka

4.1 Inner Products

Definition 4.1.1. We define the **dot product** $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ of two vectors, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ by the rule

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \boldsymbol{v} = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

Example 4.1.2. Given $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} -5 \\ 0 \\ 3 \end{bmatrix}$, their dot product is

$$\boldsymbol{u} \cdot \boldsymbol{v} = 1(-5)) + 2(0) + 3(3) = -5 + 0 + 9 = \boxed{4}.$$ ∎

Let A be an $m \times n$ matrix and let $\boldsymbol{x} \in \mathbb{R}^n$. Let $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_m \in \mathbb{R}^n$ be the row vectors of A. Thus, $A = \begin{bmatrix} \boldsymbol{r}_1 \\ \boldsymbol{r}_2 \\ \vdots \\ \boldsymbol{r}_m \end{bmatrix}$. Then we can define the matrix-vector product using dot products: $A\boldsymbol{x} = \begin{bmatrix} \boldsymbol{r}_1 \cdot \boldsymbol{x} \\ \boldsymbol{r}_2 \cdot \boldsymbol{x} \\ \vdots \\ \boldsymbol{r}_m \cdot \boldsymbol{x} \end{bmatrix}$.

There are many algebraic reasons to extend the dot product over other fields, such as $\mathbb{Z}_p$ and $\mathbb{C}$, like in the above redefinition of matrix multiplication. For another example, the dot product is used in the creation of error-correcting codes which involved vectors of $\mathbb{Z}_2^n$, but the geometric benefits of the dot product are absent on these fields and this is the very goal of this chapter. Furthermore, over complex vector spaces, we vowed to never use the matrix transpose $^\top$. Instead, we use the conjugate transpose *. The reason we made such a vow will be presented below in this section. This changes how we compute dot products of complex vectors.

Definition 4.1.3. We define the **Hermitian product**† $\cdot : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ of two vectors, $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{C}^n$ by the rule

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^* \boldsymbol{v} = \begin{bmatrix} \overline{u_1} & \overline{u_2} & \dots & \overline{u_n} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \overline{u_1} v_1 + \overline{u_2} v_2 + \ldots + \overline{u_n} v_n.$$

Example 4.1.4. Let $\boldsymbol{u} = (1+i,\ i,\ 3-i)$ and $\boldsymbol{v} = (1+i,\ 2,\ 4i)$. Find $\boldsymbol{u} \cdot \boldsymbol{v}$ and $\boldsymbol{v} \cdot \boldsymbol{u}$.

$$\boldsymbol{u} \cdot \boldsymbol{v} = (\overline{1+i})(1+i) + \bar{i}(2) + (\overline{3-i})(4i) = (1-i)(1+i) - 2i + (3+i)(4i) = 2 - 2i + 12i - 4 = \boxed{-2 + 10i}$$

$$\boldsymbol{v}\cdot\boldsymbol{u} = (\overline{1+i})(1+i) + \overline{2}(i) + (\overline{4i})(3-i) = (1-i)(1+i) + 2i - (4i)(3-i) = 2 + 2i - 12i - 4 = \boxed{-2-10i}$$

■

Note that since $\mathbb{R}^n \subseteq \mathbb{C}^n$ and $\overline{x} = x$ whenever $x \in \mathbb{R}$, the definition of the Hermitian product generalizes the notion of the dot product. But as the dot product is a possible operation on $\mathbb{C}^n$ (and there are situations where one might allow the dot product of complex vectors), we will not refer to this generalization as the dot product, but instead as the Hermitian product. Both the dot product over $\mathbb{R}^n$ and the Hermitian product over $\mathbb{C}^n$ will be denoted $\boldsymbol{u}\cdot\boldsymbol{v}$. In an attempt to unify the vocabulary here, we will call the dot product on $\mathbb{R}^n$ and the Hermitian product on $\mathbb{C}^n$ the (standard) **inner product**[††] on F^n, where F could be $\mathbb{R}$ or $\mathbb{C}$.

Theorem 4.1.5. *Let F be $\mathbb{R}$ or $\mathbb{C}$. Let $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in F^n$ and let $c \in F$. Then*

(i) $\boldsymbol{u}\cdot(\boldsymbol{v}+\boldsymbol{w}) = \boldsymbol{u}\cdot\boldsymbol{v} + \boldsymbol{u}\cdot\boldsymbol{w}$;

(ii) $\boldsymbol{u}\cdot(c\boldsymbol{v}) = c(\boldsymbol{u}\cdot\boldsymbol{v})$;

(iii) $\boldsymbol{u}\cdot\boldsymbol{v} = \boldsymbol{v}\cdot\boldsymbol{u}$;

(iv) $\boldsymbol{u}\cdot\boldsymbol{u} \geq 0$, *and* $\boldsymbol{u}\cdot\boldsymbol{u} = 0$ *if and only if* $\boldsymbol{u} = \boldsymbol{0}$.

Properties (i) and (ii) above show us that inner multiplication on the left is a linear transformation $F^n \to F$ for each fixed vector $\boldsymbol{u} \in F^n$. When $F = \mathbb{R}$, (iii) simplifies to just be $\boldsymbol{u}\cdot\boldsymbol{v} = \boldsymbol{v}\cdot\boldsymbol{u}$, that is, the dot product is *symmetric*. Combining all these three properties shows that $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w} = \boldsymbol{u}\cdot\boldsymbol{w} + \boldsymbol{v}\cdot\boldsymbol{w}$ and $(c\boldsymbol{u})\cdot\boldsymbol{v} = c(\boldsymbol{u}\cdot\boldsymbol{v})$ whenever $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \mathbb{R}^n$ and $c \in \mathbb{R}$. Thus, the dot product is linear in the first factor and linear in the second factor, that is, we say the dot product is *bilinear*. These properties derive from the properties of transposition.

Conversely, the Hermitian product is *conjugate-symmetric*, that is, the vector commute at the price of complex conjugation, as shown in (iii). As such, inner multiplication on the right is not exactly a linear transformation. Like with the real vector spaces, $(\boldsymbol{u}+\boldsymbol{v})\cdot\boldsymbol{w} = \boldsymbol{u}\cdot\boldsymbol{w} + \boldsymbol{v}\cdot\boldsymbol{w}$ still holds for all $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w} \in \mathbb{C}^n$. On the other hand, we get $(c\boldsymbol{u})\cdot\boldsymbol{v} = \overline{c}(\boldsymbol{u}\cdot\boldsymbol{v})$, that is, we can only factor scalars from the first factor if we take their conjugate. Thus, the Hermitian product is linear in the second factor and *conjugate-linear* in the second factor, that is, we say the Hermitian product is *sequilinear*. These properties derive from the properties of transposition.

Lastly, (iv) is commonly referred to as *positive-definite property*[‡], *positive* because $\boldsymbol{x}\cdot\boldsymbol{x} \geq 0$ and *definite* because $\boldsymbol{u}\cdot\boldsymbol{u} = 0$ if and only if $\boldsymbol{u} = 0$.

Definition 4.1.6. The **length** (or **norm**) of a vector $\boldsymbol{v} \in F^n$ is the nonnegative scalar $\|\boldsymbol{v}\|$ defined by

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}.$$

We say that $\boldsymbol{v}$ is a **unit vector** if $\|\boldsymbol{v}\| = 1$.

It is useful to note that $\|\boldsymbol{v}\|^2 = \boldsymbol{v}\cdot\boldsymbol{v}$.

Theorem 4.1.7. *Let $\boldsymbol{u} \in F^n$ and $c \in F$. Then*

$$\|c\boldsymbol{v}\| = |c|\|\boldsymbol{v}\|$$

Example 4.1.8. Let $\boldsymbol{v} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ -2 \end{bmatrix} \in \mathbb{R}^4$. Then the length of $\boldsymbol{v}$ is

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}} = \sqrt{1(1) + 0(0) + 2(2) - 2(-2)} = \sqrt{1+4+4} = \sqrt{9} = \boxed{3}.$$

Since the length of $\boldsymbol{v}$ is 3, $\boldsymbol{v}$ is not a unit vector. On the other hand, let $\boldsymbol{u} = \frac{1}{3}\boldsymbol{v} = \begin{bmatrix} 1/3 \\ 0 \\ 2/3 \\ -2/3 \end{bmatrix}$. Then

$$\|\boldsymbol{u}\| = \left\|\frac{1}{3}\boldsymbol{v}\right\| = \frac{1}{3}\|\boldsymbol{v}\| = \frac{1}{3}(3) = \boxed{1}.$$

Therefore, $\boldsymbol{u}$ is a unit vector in the same direction as $\boldsymbol{v}$. ■

The process of constructing a unit vector in the same direction as a given vector is called **normalization**. The normalization of any nonzero vector $\boldsymbol{v}$ is given by $\frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}$. The zero vector cannot be normalized. In particular, the zero vector does not point in any direction.

Example 4.1.9. Let $\boldsymbol{u} = (1+i,\ i,\ 3-i)$. Find $\|\boldsymbol{u}\|$.

$$\|\boldsymbol{u}\| = \sqrt{(\overline{1+i})(1+i) + (\bar{i})i + (\overline{3-i})(3-i)} = \sqrt{(1-i)(1+i) + (-i)i + (3+i)(3-i)} = \sqrt{2+1+10} = \boxed{\sqrt{13}}$$
■

Definition 4.1.10. For $\boldsymbol{u}, \boldsymbol{v} \in F^n$, the **distance** between $\boldsymbol{u}$ and $\boldsymbol{v}$, denoted as $\text{dist}(\boldsymbol{u}, \boldsymbol{v})$, is the length of the vector $\boldsymbol{u} - \boldsymbol{v}$, that is,

$$\text{dist}(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|.$$

Example 4.1.11. Let $\boldsymbol{u} = \begin{bmatrix} 7 \\ 1 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Then

$$\text{dist}(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\| = \left\| \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right\| = \sqrt{4^2 + (-1)^2} = \boxed{\sqrt{17}}.$$
■

Exercises

(Go to Solutions)

For Exercises 1-29, compute the given inner products.

1. $\begin{bmatrix} 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

♠ 2. $\left(3 \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right) \cdot \left(2 \begin{bmatrix} 2 \\ -3 \end{bmatrix} \right)$

3. $\begin{bmatrix} 7 \\ 6 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -1 \\ 5 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 7 \\ 1 \\ -5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

5. $\left(5 \begin{bmatrix} 7 \\ 1 \\ -5 \end{bmatrix} \right) \cdot \left(2 \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \right)$

♠ 6. $\left(\begin{bmatrix} 1 \\ -5 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ 2 \\ -3 \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$

♠ 7. $\left(2 \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix} \right) \cdot \left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ -5 \\ 3 \end{bmatrix} \right)$

♠ 8. $\begin{bmatrix} 2 \\ 1 \\ -3 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 7 \\ 1 \\ -1 \\ 2 \end{bmatrix}$

9. $\left(2 \begin{bmatrix} 7 \\ 1 \\ -3 \\ 2 \end{bmatrix} \right) \cdot \left(3 \begin{bmatrix} 7 \\ 1 \\ -1 \\ 2 \end{bmatrix} \right)$

♠ 10. $\left(\begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} \right) \cdot \begin{bmatrix} -1 \\ -1 \\ 3 \end{bmatrix}$

♠ 11. $\begin{bmatrix} -i \\ 1+i \end{bmatrix} \cdot \begin{bmatrix} i \\ 1 \end{bmatrix}$

♠ 12. $\begin{bmatrix} i \\ 2i \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 2-i \\ i \\ 3i \end{bmatrix}$

♠ 13. $\begin{bmatrix} i \\ 2i \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2-i \\ 4i \end{bmatrix}$

♠ 14. $\begin{bmatrix} 2-i \\ i \\ 3i \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 2-i \\ 4i \end{bmatrix}$

15. $\begin{bmatrix} 1+i \\ 2-i \\ 3+i \\ 4-i \end{bmatrix} \cdot \begin{bmatrix} 1-i \\ 2+i \\ 3-i \\ 4+i \end{bmatrix}$

16. $\left(i \begin{bmatrix} 1+i \\ 2-i \\ 3+i \\ 4-i \end{bmatrix} \right) \cdot \left(\tfrac{1}{2} \begin{bmatrix} 1-i \\ 2+i \\ 3-i \\ 4+i \end{bmatrix} \right)$

For Exercises 17-26, find the norm of the vector, that is, $\|\boldsymbol{u}\|$.

17. $\begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$

♠ 18. $\boldsymbol{u} = \begin{bmatrix} 7 \\ 1 \\ 5 \end{bmatrix}$

♠ 19. $\begin{bmatrix} 2 \\ 1 \\ -3 \\ 2 \end{bmatrix}$

20. $2 \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} - 3 \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$

21. $\begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix} + \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$

♠ 22. $\begin{bmatrix} 1 \\ -i \\ 2i \end{bmatrix}$

23. $\begin{bmatrix} i \\ 2i \\ 3 \end{bmatrix}$

24. $\begin{bmatrix} 2-i \\ i \\ 3i \end{bmatrix}$

♠ 25. $\begin{bmatrix} -1 \\ 2-i \\ 4i \end{bmatrix}$

26. $2 \begin{bmatrix} 1+i \\ 2-i \\ 3+i \\ 4-i \end{bmatrix} - \begin{bmatrix} 1-i \\ 2+i \\ 3-i \\ 4+i \end{bmatrix}$

♠ 27. Compute $\operatorname{dist}(\boldsymbol{u}, \boldsymbol{v})$, if

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 5 \\ 2 \\ 2 \end{bmatrix}$$

♠ 28. Compute $\operatorname{dist}(\boldsymbol{u} - \boldsymbol{v}, 2\boldsymbol{w} + \boldsymbol{x})$, if

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 0 \\ -5 \\ 3 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

29. Compute $\operatorname{dist}((\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}, 2\boldsymbol{x})$, if

$$\boldsymbol{u} = \begin{bmatrix} 3 \\ -1 \\ -2 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}, \boldsymbol{w} = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

30. Find x such that $\begin{bmatrix} 1 \\ 1 \\ x \end{bmatrix} \cdot \begin{bmatrix} x \\ 2 \\ -3 \end{bmatrix} = 1$.

31. For which real numbers x make $\begin{bmatrix} 1/2 \\ 1/3 \\ x \end{bmatrix}$ a unit vector?

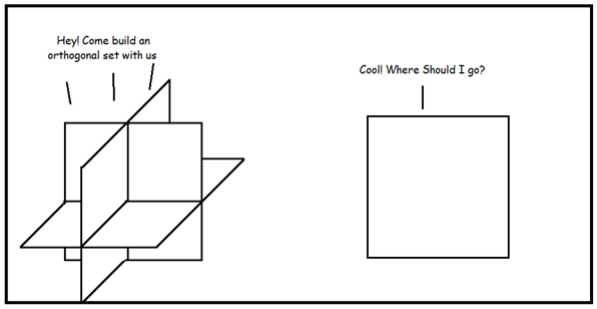
†Many textbooks alternatively define the Hermitian product as $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \overline{\boldsymbol{v}}$. Although this does not at all change the theory and applications of the Hermitian product, it does change intermediate calculations. Be cautious if comparing with other sources since there is no universal consensus.

††More generally, an *inner product* on F^n is any function $F^n \times F^n \to F$ which satisfies the axioms of Theorem 4.1.5.

‡This property is the reason we will not be considering finite fields in this chapter. The dot product over any field is always symmetric and bilinear. The positive-definite condition fails for these fields and many others. Positive-definition is needed to establish the geometric properties we seek.

"To be trusted is a greater compliment than being loved." – George MacDonald

4.2 Orthogonality



Definition 4.2.1. Two nonzero[†] vectors $\boldsymbol{u}, \boldsymbol{v} \in F^n$ are **orthogonal** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Example 4.2.2. Given the vectors $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$, and $\boldsymbol{w} = \begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$, determine if $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v}$ or $\boldsymbol{w}$.

$$\boldsymbol{u} \cdot \boldsymbol{v} = 1(1) + 2(1) + 3(-1) = 3 - 3 = 0.$$

Thus, $\boldsymbol{u}$ is orthogonal to $\boldsymbol{v}$.

$$\boldsymbol{u} \cdot \boldsymbol{w} = 1(2) + 2(3) + 3(5) = 23 \neq 0.$$

Thus, $\boldsymbol{u}$ is not orthogonal to $\boldsymbol{w}$.

$$\boldsymbol{v} \cdot \boldsymbol{w} = 1(2) + 1(3) - 1(5) = 5 - 5 = 0.$$

Thus, $\boldsymbol{v}$ is orthogonal to $\boldsymbol{w}$. ∎

Theorem 4.2.3 (Pythagorean Theorem)**.** *Two vectors $\boldsymbol{u}, \boldsymbol{v} \in F^n$ are orthogonal[††] then*

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2.$$

Proof. By assumption, $\boldsymbol{u} \cdot \boldsymbol{v} = 0 = \overline{\boldsymbol{u} \cdot \boldsymbol{v}} = \boldsymbol{v} \cdot \boldsymbol{u}$. Note that $\|\boldsymbol{u}+\boldsymbol{v}\|^2 = (\boldsymbol{u}+\boldsymbol{v}) \cdot (\boldsymbol{u}+\boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{u} + \boldsymbol{v} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{u} + 0 + 0 + \boldsymbol{v} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$. □

We call the above equation the Pythagorean Theorem because how it resembles the trigonometric equation, after all it states that a sum of squares is equal to a square under certain conditions, but there is a deeper geometry connection. First of all, if two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are linearly independent, then they span a plane in $\mathbb{R}^n$, but they also form a triangle with sides $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{u} + \boldsymbol{v}$, where the three points correspond to the tail of $\boldsymbol{u}$ (which is also the tail of $\boldsymbol{u} + \boldsymbol{v}$), the head of $\boldsymbol{u}$ (which is also the tail of $\boldsymbol{v}$), and the head of $\boldsymbol{v}$ (which is also the head of $\boldsymbol{u} + \boldsymbol{v}$). The lengths of the three sides of this triangle is $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$, and $\|\boldsymbol{u} + \boldsymbol{v}\|$. Therefore, the above theorem is saying that the sums of squares of two sides of the a triangle equal the square of the other side. This is exactly the statement of the Pythagorean Theorem from Trigonometry, but furthermore the Trigonometric Pythagorean Theorem guarantees this can only happen if there is a right angle present, that is, the angle between the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ must be a right angle. This proves the notion that two vectors are orthogonal if and only if the angle between the two vectors is $90°$.

An alternative way of expressing a plane in F^3 is by orthogonality. For example, we can construct a plane in F^3 by taking all vectors in F^3 that are orthogonal to some fixed vector $\boldsymbol{n} = (a,\ b,\ c)$, called the **normal vector** of the plane. If $\boldsymbol{x} = (x,\ y,\ z)$ is a vector on the plane, then

$$\boldsymbol{n} \cdot \boldsymbol{x} = ax + by + cz = 0.$$

This produces the equation of a plane which passes through the origin (notice that $\mathbf{0}$ is a solution here). In order to have the plane pass through the particular vector $\boldsymbol{x}_0$ we replace $\boldsymbol{x}$ with $\boldsymbol{x} - \boldsymbol{x_0}$, that is,

$$\begin{aligned} \boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{x_0}) &= 0 \\ a(x - x_0) + b(y - y_0) + c(z - z_0) &= 0 \\ ax + by + cz &= d, \end{aligned}$$

gives an equation of the plane in F^3 through the point (x_0, y_0, z_0) and orthogonal to $\boldsymbol{n}$.

Example 4.2.4. The solutions to the equation

$$3(x-4) + 5(y+5) - 2z = 0 \quad \Rightarrow \quad 3x + 5y - 2z = -13$$

is the plane containing the point $(4, -5, 0)$ and is perpendicular to $(3, 5, -2)$. ∎

Similarly, this strategy will create a hyperplane in F^n. In particular, the hyperplane containing $\boldsymbol{x}_0$ and is orthogonal to $\boldsymbol{n}$ will be the solutions $\boldsymbol{x}$ to the linear equation, given in **point-normal form**,

$$\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0. \tag{4.2.1}$$

Definition 4.2.5. A set of nonzero vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\} \subseteq F^n$ is called an **orthogonal set** if each pair of distinct vectors is orthogonal, that is, $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ whenever $i \neq j$. An orthogonal set of vectors is called **orthonormal** if each vector in the set is also a unit vector.

Of course, any orthogonal set can be transformed into an orthonormal set by normalizing each vector, that is, replacing each vector $\boldsymbol{v}$ with $\dfrac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}$.

Example 4.2.6. Let $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \boldsymbol{v}_3 = \begin{bmatrix} -5 \\ 4 \\ -1 \end{bmatrix}$. Then the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is orthogonal with respect to the dot product. To see this, we check the three dot products:

$$\begin{aligned} \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 &= 1 + 2 - 3 = 0 \\ \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 &= -5 + 8 - 3 = 0 \\ \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 &= -5 + 4 + 1 = 0. \end{aligned}$$

∎

Theorem 4.2.7. *If $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$ is an orthogonal set of nonzero vectors, then S is linearly independent.*

The requirement that an orthogonal set contain only nonzero vectors is critical. Note that $\mathbf{0} \cdot \boldsymbol{v} = 0$ for any vector $\boldsymbol{v}$. Thus, as we defined an orthogonal pair, the zero vector is orthogonal to every vector. This is a truly exceptional quality shared by no other vector. For this reason, the zero vector must be treated definitely in the theory of orthogonality and it is thus excluded from orthogonal sets. While this may seem

somewhat arbitrary, this is truly consequence driven. The main consequence is that Theorem 4.2.7 would fail if $\mathbf{0}$ were allowed in. It should also be noted that for a singleton $\{\boldsymbol{v}\}$, this set is orthogonal if and only if $\boldsymbol{v} \neq 0$, the same condition that guarantees that $\{\boldsymbol{v}\}$ is linearly independent. Notice that the condition that all pairs be orthogonal is vacuously true in this case as there are no pairs which are not orthogonal (since there are no pairs at all). This also holds for the empty set $\emptyset$. It has no pairs of distinct vectors which are not orthogonal, hence, the empty set is an orthogonal set.

An orthogonal (orthonormal) spanning set of F^n is necessarily a basis, called an **orthogonal (orthonormal) basis**. For example, the standard basis in F^n is an orthonormal basis.

Theorem 4.2.8. *Let W be a subspace of F^n. Let $W^\perp = \{\boldsymbol{x} \in F^n \mid \boldsymbol{w} \cdot \boldsymbol{x} = 0,\ \forall \boldsymbol{w} \in W\}$, called the **orthogonal complement** of W. Then $W^\perp$ is also a subspace of F^n.*

Example 4.2.9. Let W be the z-axis in $\mathbb{R}^3$. Then $W^\perp$ is the xy-plane. ∎

Theorem 4.2.10. *Let A be an $m \times n$ matrix. Then*

$$(\operatorname{Row} A)^\perp = \operatorname{Nul} A.^\ddagger$$

Example 4.2.11. Let W be the subspace of $\mathbb{R}^6$ spanned by the vectors

$$\boldsymbol{w}_1 = (1, 3, -2, 0, 2, 0),\ \boldsymbol{w}_2 = (2, 6, -5, -2, 4, -3),\ \boldsymbol{w}_3 = (0, 0, 5, 10, 0, 15),\ \boldsymbol{w}_4 = (2, 6, 0, 8, 4, 18).$$

Construct a base for $W^\perp$.

We can arrange these vectors into a matrix A:

$$A = \begin{bmatrix} 1 & 3 & -2 & 0 & 2 & 0 \\ 2 & 6 & -5 & -2 & 4 & -3 \\ 0 & 0 & 5 & 10 & 0 & 15 \\ 2 & 6 & 0 & 9 & 4 & 18 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 0 & 4 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

such that $W = \operatorname{Row}(A)$. By the previous result, $W^\perp = \operatorname{Row}(A)^\perp = \operatorname{Nul}(A)$. As we know how to find a basis for a null space given the echelon form of the matrix,

$$W^\perp = \operatorname{Span}\{(-3, 1, 0, 0, 0, 0),\ (-4, 0, -2, 1, 0, 0),\ (-2, 0, 0, 0, 1, 0)\}.$$ ∎

Despite our reluctance to use any fields other than $\mathbb{R}$ and $\mathbb{C}$, the entirety of this section[‡‡] is transferable to any field using the dot product or Hermitian product.

Exercises

(Go to Solutions)

♠ 1. Find k so that $\begin{bmatrix} k \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ is orthogonal.

For Exercises 2-5, determine whether the two given vectors are orthogonal or not. If not, adjust a value in one of the vectors to make them orthogonal. Answers may vary.

2. $\begin{bmatrix} 4 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ -8 \end{bmatrix}$

3. $\begin{bmatrix} 6 \\ 1 \end{bmatrix}, \begin{bmatrix} 8 \\ -24 \end{bmatrix}$

4. $\begin{bmatrix} -3 \\ -5 \end{bmatrix}, \begin{bmatrix} 5 \\ -3 \end{bmatrix}$

5. $\begin{bmatrix} 2 \\ 8 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \end{bmatrix}$

For Exercises 6-9, find a hyperplane in F^n containing the vector $\boldsymbol{x}_0$, listed first, and whose normal vector is $\boldsymbol{n}$, listed second.

♠ 6. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 0 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 2 \\ -1 \end{bmatrix}$

♠ 8. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} -3 \\ -6 \\ 3 \\ 2 \\ -5 \end{bmatrix}$

9. $\begin{bmatrix} 10 \\ -8 \\ 6 \\ -4 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \\ 5 \\ 7 \\ 9 \end{bmatrix}$

For Exercises 10-17, find a basis for the orthogonal complement of the subspace W, given below. Answers may vary.

10. $\text{Span}\left\{ \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$

♠ 11. $\text{Span}\left\{ \begin{bmatrix} 3 \\ -1 \\ 9 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix} \right\}$

♠ 12. $\text{Span}\left\{ \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} \right\}$

♠ 13. $\text{Span}\left\{ \begin{bmatrix} 3 \\ 3 \\ -12 \\ 6 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \\ 12 \\ -4 \end{bmatrix} \right\}$

♠ 14. $\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| 2x_1 + 3x_2 = 0 \right\}$

15. $\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \middle| x_1 + x_2 = 0 \right\}$

♠ 16. $\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \middle| \, 3x_1 - x_2 + 2x_3 = 0 \right\}$

♠ 17. $\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \middle| \, x_1 + x_2 = 0, x_1 - x_3 = 0 \right\}$

For Exercises 18-18, find a hyperplane in F^n containing the given vectors $\boldsymbol{x}_0, \boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$. [**Hint**: If $W = \text{Span}\{\boldsymbol{x}_0 - \boldsymbol{x}_1, \boldsymbol{x}_0 - \boldsymbol{x}_2, \ldots\}$, find a normal vector by computing a basis for $W^\perp$.]

18. $(1, 2, 3)$, $(-1, 4, 1)$, $(4, -2, 8)$, $(0, 1, 6)$

---

†When it comes to defining orthogonality, we have to exclude the zero vector. Note that $\mathbf{0} \cdot \boldsymbol{v} = 0$ for every vector $\boldsymbol{v}$. As such, if we allowed the zero vector to be orthogonal then it would be orthogonal to every vector, including itself, which would provide needless counterexamples to theorems and geometric interpretations that will follow.

††For real vector spaces, the Pythagorean Theorem becomes an "if and only if" statement, that is, *Two vectors* $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ *are orthogonal if and only if*

$$\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2.$$

The proof of the other direction is based upon the observation that the equality of squared norms holds only if $\boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{v} \cdot \boldsymbol{u} = 0$. In the real case, of course, the left-hand side simplifies to $2(\boldsymbol{u} \cdot \boldsymbol{v})$, which then implies that $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. In the complex case, counterexamples exist. For example, let $\boldsymbol{u} = (1, 1, 1, 1)$ and $\boldsymbol{v} = (0, 0, 0, i)$. Note that $\|\boldsymbol{u}\| = 2$ and $\|\boldsymbol{v}\| = 1$. Hence, $\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 = 5$. On the other hand, $\|\boldsymbol{u} + \boldsymbol{v}\|^2 = 1 + 1 + 1 + 2 = 5$, but $\boldsymbol{u} \cdot \boldsymbol{v} = i \neq 0$. The defect in extending the proof to all complex vectors here is that equality will hold whenever the Hermitian product between the two vectors is purely imaginary. Of course, the only purely imaginary number which is also real is zero, hence why it holds in the real case.

†For real vector spaces, recall that $\text{Row}(A) = \text{Col}(A^\top)$. With the perspective, Theorem 4.2.10 tells us that $\text{Nul}(A)^\perp = \text{Col}(A^\top)$. Hence, the orthogonal complement $\perp$ and transpose $\top$ are, in same way, dual notions of each other.

‡‡There is one important partial exception to this claim, namely the Pythagorean Theorem. We say this as a partial exception because while the algebraic statement and proof still remain valid (note that $\|\boldsymbol{u}\|^2 = \boldsymbol{u} \cdot \boldsymbol{u}$ and is valid for any vector space which we desire to introduce the dot product), the geometric interpretations about the norm (if it is defined via the dot product) do not necessarily transfer to general vector spaces.

"We can never obtain peace in the outer world until we make peace with ourselves." – Dalai Lama

4.3 Outer Products

We have already discussed some special families of matrices, e.g. nonsingular, elementary, diagonal, and triangular matrices. In this section, we will explore some more special types of matrices and some of their properties. These matrices are important in our study of orthogonality and later when we study eigenvalues.

Definition 4.3.1. A real matrix A is **symmetric** if $A^\top = A$. A complex matrix A is **Hermitian** if $A^* = A$.

Example 4.3.2. Let $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & i & 1+i \\ -i & -5 & 2-i \\ 1-i & 2+i & 3 \end{bmatrix}$. Since $A^\top = A$ and $B^* = B$ (convince yourself!), A is a symmetric matrix and B is an Hermitian matrix. ∎

The theory of symmetric and Hermitian matrices is almost identical. We will primarily focus on real matrices below and will only specify Hermitian matrices when a critical difference arises.

Theorem 4.3.3. *If A and B are $n \times n$ symmetric (Hermitian) matrices and $r \in \mathbb{R}$ ($r \in \mathbb{C}$), then*

(i) $A + B$ is symmetric (Hermitian)

(ii) rA is symmetric (Hermitian)

(iii) AB is symmetric (Hermitian) if and only if $AB = BA$

(iv) if A is invertible, then A^{-1} is symmetric (Hermitian).

Example 4.3.4. Note that $A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 3 \\ 3 & 1 \end{bmatrix}$ are symmetric matrices. Likewise,

$$A + B = \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} + \begin{bmatrix} 5 & 3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+3 \\ 2+3 & 0+1 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 5 & 1 \end{bmatrix}$$

is likewise symmetric, as well as

$$5A = 5 \begin{bmatrix} 1 & 2 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 0 \end{bmatrix}.$$ ∎

Theorem 4.3.5. *For any matrix A, the matrices $A^\top A$ and $AA^\top$ are symmetric (A^*A and AA^* are Hermitian).*

Example 4.3.6. Let $A = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix}$. Then

$$A^\top A = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 10 & -2 & -11 \\ -2 & 4 & -8 \\ -11 & -8 & 41 \end{bmatrix}$$

and

$$AA^\top = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} 21 & -17 \\ -17 & 34 \end{bmatrix}. \qquad \blacksquare$$

Definition 4.3.7. Let $P : F^n \to F^n$ be a linear transformation. We say that P is a **projection** if $P \circ P = P$.

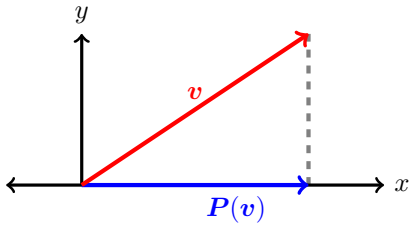
Let A be the standard matrix of P. Then the property that $P \circ P = P$ translate to mean that $A^2 = A$. Any matrix (necessarily a square) which satisfies this identity is called an **idempotent** matrix and is necessarily the standard matrix of a projection.

Let $\boldsymbol{x} \in F^n$. Then $\boldsymbol{y} = P(\boldsymbol{x})$ is an arbitrary element of the range of P. By definition, we have that

$$P(\boldsymbol{y}) = P(P(\boldsymbol{x})) = P \circ P(\boldsymbol{x}) = P(\boldsymbol{x}) = \boldsymbol{y},$$

that is, a projection is exactly a linear transformation that fixes its image. Essentially this means that while some of the coordinates of $\boldsymbol{x}$ are unaltered the other coordinates are forgotten.

Geometrically, a projection is a map from the ambient space which projects onto some subspace (the range of the projection). In the process of projecting into the subspace, some information about the vectors is removed so that they can fit inside the smaller space. The image of a vector $\boldsymbol{v}$ is the shadow cast by $\boldsymbol{v}$ onto the subspace associated with the projection P.



Example 4.3.8. For example, the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

are easily seen to be idempotent matrices and hence correspond to projections. The first matrix is the projection in $\mathbb{R}^2$ onto the x-axis where the y-coordinate is forgotten and replaced with 0. Any point already on the x-axis has the form $(x, 0)$ and is unaffected by the projection.

The second matrix is a projection in $\mathbb{R}^3$ onto the y-axis, where the x- and z-coordinates are discarded. Likewise, the third matrix is a projection in $\mathbb{R}^3$ onto the xy-plane, where the z-coordinate is the only information forgotten. In either case, elements already in the subspace are not altered by the projection. $\blacksquare$

Example 4.3.9. The matrices

$$\begin{bmatrix} 2 & 2 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} -8 & 4 & 1 \\ -18 & 9 & 2 \\ 0 & 0 & 1 \end{bmatrix}$$

are also idempotent matrices. (Convince yourself of this). The first matrix corresponds to projection of $\mathbb{R}^2$ onto the line spanned by $(2, -1)$, that is, the line $y = -\frac{1}{2}x$. The second matrix corresponds to projection of $\mathbb{R}^3$ onto the plane spanned by $(4, 9, 0)$ and $(1, 2, 1)$, that is, the plane $9x - 4y - z = 0$. $\blacksquare$

If A is idempotent, then multiplication by A is projection onto $\text{Col}(A)$.

Definition 4.3.10. Let $\boldsymbol{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} \in F^n$. Then the **outer product** (or **matrix product**) of $\boldsymbol{u}$ and $\boldsymbol{v}$, denoted $\boldsymbol{u} \otimes \boldsymbol{v}$, is the $n \times n$ matrix of the form $\begin{bmatrix} u_i v_j \end{bmatrix}$.

More compactly, we have that $\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^\top$ (or $\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^*$ for complex vectors). This highly resembles the definition of the inner product $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^\top \boldsymbol{v}$ (or $\boldsymbol{u}^* \boldsymbol{v}$ for complex vectors). Despite this similarity, the outer product is a matrix and the inner product is a scalar. Neither of them is a vector. Of course, these two products are interconnected, justifying the complementary names, by the following formula:

$$(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{w} = (\boldsymbol{v} \cdot \boldsymbol{w})\boldsymbol{u}.$$

Proof.

$$(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{w} = (\boldsymbol{u}\boldsymbol{v}^\top)\boldsymbol{w} = \boldsymbol{u}(\boldsymbol{v}^\top \boldsymbol{w}) = \boldsymbol{u}(\boldsymbol{v} \cdot \boldsymbol{w}) = (\boldsymbol{v} \cdot \boldsymbol{w})\boldsymbol{u}. \qquad \square$$

Essentially, the adjectives inner versus outer can be a mnemonic to describe the location of the $^\top$ (or *).

Example 4.3.11. Let $\boldsymbol{u} = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 2 \\ -2 \\ 3 \\ -6 \end{bmatrix}$. Then $\boldsymbol{u} \otimes \boldsymbol{v} = \begin{bmatrix} 2 & -2 & 3 & -6 \\ -4 & 4 & -6 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. ■

Theorem 4.3.12. *Let $\boldsymbol{u}$ be a unit vector. Then $A = \boldsymbol{u} \otimes \boldsymbol{u}$ is an idempotent matrix, and the matrix transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is a projection onto the subspace* $\text{Span}\{\boldsymbol{u}\}$.

Example 4.3.13. Consider the vector $\boldsymbol{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Its normalization is $\boldsymbol{u} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$. Then the matrix $A = \boldsymbol{u} \otimes \boldsymbol{u} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$ is an idempotent matrix since

$$A^2 = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/4+1/4 & 1/4+1/4 \\ 1/4+1/4 & 1/4+1/4 \end{bmatrix} = A.$$

The mapping $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is the projection of a vector in $\mathbb{R}^2$ onto the line $y = x$. ■

Definition 4.3.14. Let A be an $n \times n$ matrix. We say that A is **nilpotent** if $A^n = 0$, the zero matrix.

Of course, it is possible for a nilpotent matrix A that $A^m = 0$ for some integer smaller than n. For example, the outer product can also be used to create nilpotent matrices such that $A^2 = 0$ for any n.

Theorem 4.3.15. *Let $\boldsymbol{u}, \boldsymbol{v} \in F^n$ such that $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal. Then $A = \boldsymbol{u} \otimes \boldsymbol{v}$ is a nilpotent matrix. In particular, $A^2 = 0$.*

Example 4.3.16. Note that $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Then $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. Then $A = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$ is nilpotent since

$$A^2 = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} = \begin{bmatrix} 4-4 & -2+2 \\ 8-8 & -4+4 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \qquad \blacksquare$$

Example 4.3.17. We note that all strictly triangular matrices are nilpotent. For example, if $A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$, then $A^2 = 0$, and hence A is nilpotent.

Likewise, if $B = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}$, then $B^2 = \begin{bmatrix} 0 & 0 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $B^3 = 0$. Hence, we have construct a nilpotent matrix such that $B^2 \neq 0$. Hence, B is a nilpotent matrix that cannot be factored as an outer product. You will notice that in the matrix B, while we had only zeros along the diagonal, we did have non-zero entries in the slant right above the main diagonal and the slant right above that one, the so called *second upper diagonal* and *third upper diagonal*. When we squared B, we gained zeros along the second diagonal but still had nonzero entries on the third diagonal. Only when we cubed the matrix did we get zeros everywhere. This is essentially how multiplication of strictly triangular matrices work. Every time we take another power, the matrix loses one more of its diagonals. Since this will eventual terminate, strictly triangular matrices are nilpotent. $\blacksquare$

Exercises

(Go to Solutions)

For Exercises 1-6, finish the matrix so that it is symmetric or Hermitian.

1. $\begin{bmatrix} * & 7 \\ * & * \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1 & 2 & * \\ * & 1 & 2 \\ 3 & * & 1 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 7 & 8 & 9 \\ * & 5 & 6 \\ * & * & 3 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 56 & * & 3 & 23 \\ 21 & 91 & 85 & * \\ * & * & 43 & 35 \\ * & 75 & * & 62 \end{bmatrix}$

5. $\begin{bmatrix} * & 6 & * & * & * \\ * & 7 & 23 & 14 & 8 \\ 8 & * & 11 & * & * \\ 9 & * & 16 & 54 & 22 \\ 13 & * & 19 & * & 72 \end{bmatrix}$

6. $\begin{bmatrix} 1 & * & 5-2i \\ -7i & 3 & * \\ * & 9i & 2 \end{bmatrix}$

For Exercises 7-9, give an example of a non-symmetric real matrix for each listed size. (Answers may vary).

7. 2×2

8. 3×3

9. 4×4

For Exercises 10-13, determine whether the matrix is Hermitian or not.

♠ 10. $\begin{bmatrix} 2 & 1-3i \\ 1+3i & -3 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 1 & 3 \\ 3 & 0 \end{bmatrix}$

♠ 12. $\begin{bmatrix} 0 & 1-i \\ 1+i & i \end{bmatrix}$

♠ 13. $\begin{bmatrix} 3 & -1 \\ -1 & 2 \end{bmatrix}$

We say a matrix A is **skew-symmetric** (or **alternating**) if $A^\top = -A$. For Exercises 14-15, finish the matrix so that it is skew-symmetric.

14. $\begin{bmatrix} * & -2 \\ * & * \end{bmatrix}$

15. $\begin{bmatrix} * & 3 & 9 & * \\ * & * & * & * \\ * & -5 & 0 & * \\ -3 & 2 & 9 & 0 \end{bmatrix}$

For Exercises 16-18, compute the given quantity.

♠ 16. Compute $\boldsymbol{u} \otimes \boldsymbol{v}$, if $\boldsymbol{u} = \begin{bmatrix} -3 \\ 2 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 4 \\ 0 \end{bmatrix}$

♠ 17. Compute $\boldsymbol{u} \otimes \boldsymbol{v}$, if $\boldsymbol{u} = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 4 \\ -1 \\ 2 \end{bmatrix}$

♠ 18. Compute $\boldsymbol{u} \otimes \boldsymbol{v}$, if $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$

For Exercises 19-22, find vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ such that their outer product is equal to the given matrix, that is, find an *outer factorization*.

19. $\boldsymbol{u} \otimes \boldsymbol{v} = \begin{bmatrix} -28 & 0 & 7 \\ -8 & 0 & 2 \\ -4 & 0 & 1 \end{bmatrix}$

20. $\boldsymbol{u} \otimes \boldsymbol{v} = \begin{bmatrix} 14 & -4 \\ -7 & 2 \\ -28 & 8 \\ 7 & -2 \end{bmatrix}$

21. $\boldsymbol{u} \otimes \boldsymbol{v} = \begin{bmatrix} 6 & 9 & 6 & 12 \\ 6 & 9 & 6 & 12 \\ 12 & 18 & 12 & 24 \\ 18 & 27 & 18 & 36 \\ 8 & 12 & 8 & 16 \end{bmatrix}$

22. $\boldsymbol{u} \otimes \boldsymbol{v} = \begin{bmatrix} 18 & 0 & 0 & 45 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -16 & 0 & 0 & -40 \end{bmatrix}$

For Exercises 23-25, determine whether the matrix is idempotent, nilpotent, or neither.

♠ 23. $\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$

♠ 24. $\begin{bmatrix} 1 & 1 & 2 \\ 1 & 0 & -3 \\ 2 & 2 & 0 \end{bmatrix}$

♠ 25. $\begin{bmatrix} 0 & 4 & -2 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix}$

♠ 26. Using the projection in Example 4.3.13, draw the image of the unit square under this projection and describe its "area."

27. Let A be a square matrix.

    (a) Show that $A + A^\top$ is a symmetric matrix.

    (b) Show that $A - A^\top$ is a skew-symmetric matrix (see Exercise 14).

    (c) Show that there exists matrices S and T such that S is symmetric, T is skew-symmetric, and $A = S + T$.

28. Let A be a square matrix. Show that if A is both idempotent and invertible then A is an identity matrix. Can a nilpotent matrix be invertible? Explain why or why not.

29. In Exercise 14, the notion of a skew-symmetric matrix was introduced. What would be the analogous definition of skew-Hermitian matrix? Provide an example of a non-real skew-Hermitian matrix.

"It's always good to take an orthogonal view of something. It develops ideas." – Ken Thompson

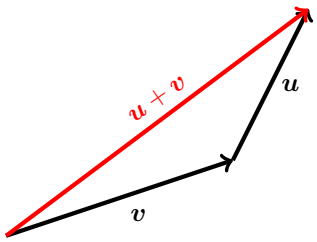
4.4 Affine Transformations

Theorem 4.4.1 (Cauchy-Schwarz Inequality)**.** *For all* $\boldsymbol{u}, \boldsymbol{v} \in F^n$*,*

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|.$$

Theorem 4.4.2 (The Triangle Inequality)**.** *For all* $\boldsymbol{u}, \boldsymbol{v} \in F^n$*,*

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|.$$



Theorem 4.4.3 (The Law of Cosines)**.** *Let* $\boldsymbol{u}, \boldsymbol{y} \in F^n$*. Then*

$$\boldsymbol{u} \cdot \boldsymbol{y} = \|\boldsymbol{u}\|\|\boldsymbol{y}\| \cos\theta,^{\dagger}$$

where θ *is the angle between the two line segments from the origin to the points identified with* $\boldsymbol{u}$ *and* $\boldsymbol{y}$*.*

Example 4.4.4. Compute the angle θ between $\boldsymbol{u} = (6, -1)$ and $\boldsymbol{v} = (1, 4)$ in $\mathbb{R}^2$.

$$\theta = \cos^{-1}\left(\frac{6(1) - 1(4)}{\sqrt{6^2 + (-1)^2}\sqrt{1^2 + 4^2}}\right) = \cos^{-1}\left(\frac{6 - 4}{\sqrt{36 + 1}\sqrt{1 + 16}}\right) = \cos^{-1}\left(\frac{2}{\sqrt{37}\sqrt{17}}\right) \approx \boxed{85.43^\circ} \quad \blacksquare$$

From the previous theorem, if the angle between two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is 90°, then the inner product $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Definition 4.4.5. A real square matrix U is called **orthogonal** if $U^\top U = I$, that is, $U^T = U^{-1}$. A complex square matrix U is said to be **unitary** if $U^* = U^{-1}$.

Theorem 4.4.6. *A square matrix* U *is orthogonal (unitary) if and only if its column vectors form an orthonormal set.*

Since $U^\top = U^{-1}$ and $UU^\top = UU^{-1} = I$, it also follows that the row vectors of an orthogonal matrix U must also form an orthonormal set.

Rotation and reflection matrices are examples of orthogonal matrices.

Example 4.4.7. The matrix

$$U = \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

is an orthogonal matrix. Note that

$$U^\top U = \begin{bmatrix} 3/\sqrt{11} & 1/\sqrt{11} & 1/\sqrt{11} \\ -1/\sqrt{6} & 2/\sqrt{6} & 1/\sqrt{6} \\ -1/\sqrt{66} & -4/\sqrt{66} & 7/\sqrt{66} \end{bmatrix} \begin{bmatrix} 3/\sqrt{11} & -1/\sqrt{6} & -1/\sqrt{66} \\ 1/\sqrt{11} & 2/\sqrt{6} & -4/\sqrt{66} \\ 1/\sqrt{11} & 1/\sqrt{6} & 7/\sqrt{66} \end{bmatrix}$$

$$= \begin{bmatrix} (9+1+1)/11 & (-3+2+1)/\sqrt{66} & (-3-4+7)/\sqrt{726} \\ (-3+2+1)/\sqrt{66} & (1+4+1)/6 & (1-8+7)/\sqrt{396} \\ (-3-4+7)/\sqrt{726} & (1-8+7)/\sqrt{396} & (1+16+49)/66 \end{bmatrix} = I_3$$ ■

Example 4.4.8. Let $U = \begin{bmatrix} \frac{1}{2}(1+i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1+i) \end{bmatrix}$. Note that

$$UU^* = \begin{bmatrix} \frac{1}{2}(1+i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1+i) \end{bmatrix} \begin{bmatrix} \frac{1}{2}(1-i) & \frac{1}{2}(1+i) \\ \frac{1}{2}(1-i) & \frac{1}{2}(-1-i) \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{1}{2} & \frac{i}{2}-\frac{i}{2} \\ -\frac{i}{2}+\frac{i}{2} & \frac{1}{2}+\frac{1}{2} \end{bmatrix} = I_2$$

Therefore, U is unitary. ■

Theorem 4.4.9. *Let U be an orthogonal (unitary) matrix, and let $\boldsymbol{x}, \boldsymbol{y} \in F^n$. Then*

$$(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y},$$

that is, the matrix transformation $\boldsymbol{x} \mapsto \boldsymbol{U}\boldsymbol{x}$ preserves inner products.

Proof.

$$(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = (U\boldsymbol{x})^\top (U\boldsymbol{y}) = (\boldsymbol{x}^\top U^\top)(U\boldsymbol{y}) = \boldsymbol{x}^\top (U^\top U)\boldsymbol{y} = \boldsymbol{x}^\top \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}.$$ □

Since multiplication by an orthogonal (unitary) matrix U preserves inner products, as a consequence, it also preserves lengths, distances, angles, and orthogonality of vectors. For example, $\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$ and $\operatorname{dist}(U\boldsymbol{x}, U\boldsymbol{y}) = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{y})$ for any vectors $\boldsymbol{x}, \boldsymbol{y}$ and any orthogonal matrix U.

Theorem 4.4.10. *Let $\mathcal{B}$ and $\mathcal{C}$ be orthonormal bases of F^n. Then the change-of-basis matrix $\underset{\mathcal{C}\leftarrow\mathcal{B}}{\mathcal{P}}$ is orthogonal (unitary).*

Definition 4.4.11. An **isometry** (or a **rigid motion**) is a function $R : F^n \to F^n$ such that $\operatorname{dist}(T(\boldsymbol{x}), T(\boldsymbol{y})) = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{y})$ for all $\boldsymbol{x}, \boldsymbol{y} \in F^n$.

Isometry, which translates as "same-measure" from Greek, are those maps that preserve distances and lengths. Using trigonometry, necessarily angles are preserved too. Therefore, an isometry is exactly a map which "moves" objects is the vector space without "distorting" them, that is, the image of a shape will be congruent to the original shape. Linear isometries are exactly multiplication by an orthogonal (unitary) matrix.

Orthogonal transformations are not the only way to make an isometry. Let $\boldsymbol{b} \in F^n$ be any vector in the vector space. We can define a transformation $T : F^n \to F^n$ by the rule $T(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{b}$, called a **translation** by $\boldsymbol{b}$. If $\boldsymbol{b} \neq \boldsymbol{0}$, then translation is not a linear map since $T(\boldsymbol{0}) = \boldsymbol{b} \neq 0$. On the other hand, $\operatorname{dist}(T(\boldsymbol{x}), T(\boldsymbol{y})) = \|(\boldsymbol{x}+\boldsymbol{b}) - (\boldsymbol{y}+\boldsymbol{b})\| = \|\boldsymbol{x}+\boldsymbol{b}-\boldsymbol{y}-\boldsymbol{b}\| = \|\boldsymbol{x}-\boldsymbol{y}\| = \operatorname{dist}(\boldsymbol{x}, \boldsymbol{y})$. If we allow for translations in a vector space, we must broaden our type of transformations to affine transformations.

Definition 4.4.12. Let $T : F^n \to F^m$ be a function. We say that T is an **affine transformation** if there exists a matrix $A \in F^{m\times n}$ and a vector $\boldsymbol{b} \in F^m$ such that $T(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ for all $\boldsymbol{x} \in F^n$.

Of course, a linear transformation is affine (when $\boldsymbol{b} = \boldsymbol{0}$) and every translation is affine (when $A = I_n$). Affine transformations are very important in linear geometry. As translation maps are ALWAYS one-to-one (onto), an affine transformation $\boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ is one-to-one (onto) if and only if the matrix transformation $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is one-to-one (onto).

Example 4.4.13. Consider the affine transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ associated to the matrix $A = \begin{bmatrix} 3 & -1 & 1 \\ -3 & 2 & 0 \\ 6 & -3 & 2 \end{bmatrix}$ and the translation vector $\boldsymbol{b} = (1, 2, 3)$. Then

$$T\left(\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 3 & -1 & 1 \\ -3 & 2 & 0 \\ 6 & -3 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ 10 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 13 \end{bmatrix}.$$

Is $\boldsymbol{y} = (2, -2, 4) \in \text{im}(T)$? To answer this question, we solve the matrix equation $A\boldsymbol{x} + \boldsymbol{b} = \boldsymbol{y} \Rightarrow A\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{b} = (1, -4, 1)$:

$$\left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ -3 & 2 & 0 & -4 \\ 6 & -3 & 2 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -4 \end{array}\right].$$

Therefore, $T(2, 1, -4) = (2, -2, 4)$. ∎

As mentioned above, an affine transformation $T : F^n \to F^n : T(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$ cannot be linear if $\boldsymbol{b} \neq 0$ on F^n. On the other hand, we can realize affine transformations as matrix multiplication in a higher dimension. After all, F^n is a naturally is a subset of F^{n+1} and consists of all those vectors in F^{n+1} whose last coordinate is zero. We may identify F^n in F^{n+1} with the hyperplane associated to the linear equation $x_{n+1} = 1$. In other words, we will translate F^n by the vector $\boldsymbol{e}_{n+1}$ and consider this n-flat. It will consist of all vectors in F^{n+1} whose last coordinate is 1. Suppose $\boldsymbol{y} = T(\boldsymbol{x}) = A\boldsymbol{x} + \boldsymbol{b}$, which is the image of $\boldsymbol{x}$ under the affine transformation. Then

$$\left[\begin{array}{c|c} A & \boldsymbol{b} \\ 0 \ldots 0 & 1 \end{array}\right]\begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} = \begin{bmatrix} A\boldsymbol{x} + \boldsymbol{b} \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ 1 \end{bmatrix}.$$

Hence, the affine transformation $T : F^n \to F^m$ can be extended to a linear transformation $L : F^{n+1} \to F^{m+1}$ in a higher dimension.

Example 4.4.14. Using the affine transformation from Example 4.4.13, the standard matrix of the affine transformation is:

$$[T] = \left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ -3 & 2 & 0 & 2 \\ 6 & -3 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right]. \quad \text{Note that } T(2, 0, -1) = \left[\begin{array}{ccc|c} 3 & -1 & 1 & 1 \\ -3 & 2 & 0 & 2 \\ 6 & -3 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{array}\right]\begin{bmatrix} 2 \\ 0 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ -4 \\ 13 \\ 1 \end{bmatrix}. \quad ∎$$

Theorem 4.4.15 (Mazur-Ulam Theorem)**.** *If $T : F^n \to F^n$ is an isometry then, for all $\boldsymbol{x} \in F^n$, $T(\boldsymbol{x}) = U\boldsymbol{x} + \boldsymbol{b}$ for some translation vector $\boldsymbol{b} \in F^n$ and orthogonal (unitary) matrix $U \in F^{n\times n}$.*

Over $\mathbb{R}^2$ there are four types of isometries: rotation around a point in the plane, reflection across a line in the plane, translation, and glide reflection (the composition of a reflection and a translation, like footprints in the sand).

Exercises

(Go to Solutions)

For Exercises 1-3, find the angle between the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathbb{R}^n$.

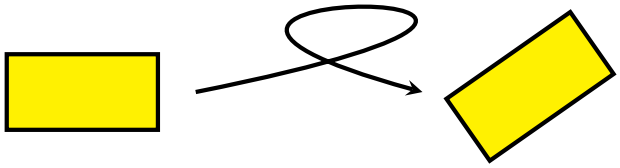
♠ 1. $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$

♠ 2. $\boldsymbol{u} = \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

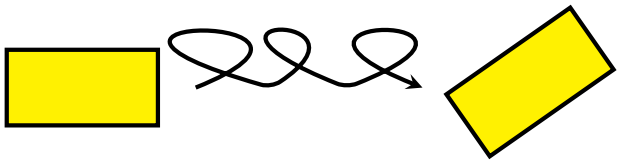
♠ 3. $\boldsymbol{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

For Exercises 4-11, identify whether of the transformation is an isometry. The graphic on the left will denote the original shape and the graphic on the right is the transformed shape.

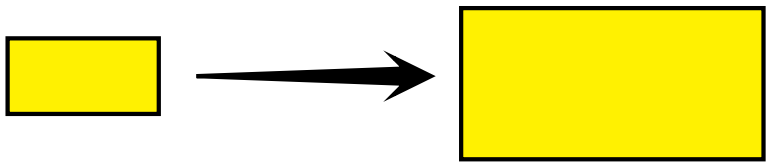
♠ 4. The shape was translated and rotated.



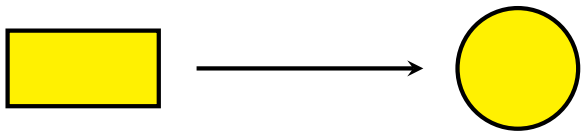
♠ 5. The shape was translated and rotated.††



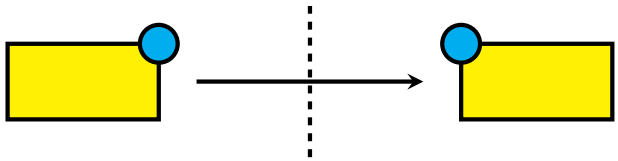
♠ 6. The shape was translated and increased in size.‡



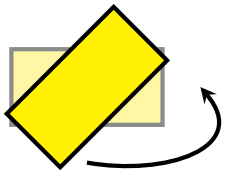
♠ 7. The object was translated and changed its shape from a rectangle to a circle, although the area stayed the same.



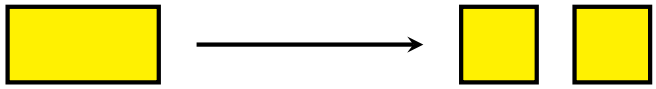
♠ 8. The object was reflected.



♠ 9. In the following transformation, the shadow object denotes the original object and the opaque object denotes the transformed object.‡‡



♠ 10. The shape was translated and has been cut.



♠ 11. The following transformation is a little more difficult to visualize. Imagine the shape is transformed by placing it back exactly the way it was before, that is, the starting and stopping shape is identical.

For Exercises 12-15, let $T : \mathbb{R}^2 \to \mathbb{R}^2 : \boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{b}$ be translation by $\boldsymbol{b}$, provided below. Let $P = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Draw P and $T(P)$ on the same gridlines.

♠ 12. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ♠ 13. $\begin{bmatrix} 0 \\ -2 \end{bmatrix}$ ♠ 14. $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ ♠ 15. $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$

For Exercises 16-18, consider the affine transformation $T : \mathbb{R}^3 \to \mathbb{R}^3 : \boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ associated to the matrix $A = \begin{bmatrix} 3 & -1 & 1 \\ -3 & 2 & 0 \\ 6 & -3 & 2 \end{bmatrix}$ and the translation vector $\boldsymbol{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

16. Compute $T(3, 1, -5)$.

17. Find a vector $\boldsymbol{x} \in \mathbb{R}^3$ such that $T(\boldsymbol{x}) = (6, 1, 3)$.

18. Is T one-to-one? Onto? Why or why not?

For Exercises 19-21, consider the affine transformation $T : \mathbb{R}^3 \to \mathbb{R}^3 : \boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ associated to the matrix

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 2 & 2 \\ 4 & 1 & -3 \end{bmatrix}$$ and the translation vector

$$\boldsymbol{b} = \begin{bmatrix} -5 \\ 6 \\ 8 \end{bmatrix}.$$

♠ 19. Compute $T(5, -2, 4)$.

♠ 20. Find a vector $\boldsymbol{x} \in \mathbb{R}^3$ such that $T(\boldsymbol{x}) = (-6, 8, 12)$.

♠ 21. Is T one-to-one? Onto? Why or why not?

For Exercises 22-24, consider the affine transformation $T : \mathbb{R}^3 \to \mathbb{R}^3 : \boldsymbol{x} \mapsto A\boldsymbol{x} + \boldsymbol{b}$ associated to the matrix

$$A = \begin{bmatrix} 6 & 4 & 6 \\ 2 & 2 & 2 \\ 10 & 8 & 6 \end{bmatrix}$$ and the translation vector

$$\boldsymbol{b} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

22. Compute $T(1, 2, 1)$.

23. Find a vector $\boldsymbol{x} \in \mathbb{R}^3$ such that $T(\boldsymbol{x}) = (7, 6, 5)$.

24. Is T one-to-one? Onto? Why or why not?

†For complex vectors, the real valued function $\cos : \mathbb{R} \to \mathbb{R}$ will need to be extended to the complex plane $\cos : \mathbb{C} \to \mathbb{C}$.

††Although the precious path of rotation and translation is different from the previous exercise, the starting and stopping positions of the shape are exactly the same. In this case, we say these two transformations **equal**. In particular, a transformation is determined by the image itself of the shape and not be the process which produces the image.

‡This is an example of a *dilation*. To shrink in size is called a *contraction*.

‡‡Notice that the center of the rectangle is not moved by the motion, that is, the transformation assigns this point to itself. We say that a point P is a *fixed point* if $P' = P$. The *identity motion* is characterized as the transformation which assigns every point to itself, that is, all points are fixed. Can you find the identity motion in this homework set?

"To raise new questions, new possibilities, to regard old problems from a new angle, requires creative imagination and marks real advance in science." – Albert Einstein

4.5 Orthogonal Projections

Theorem 4.5.1 (Parseval's Identity)**.** *Let $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$ be an orthogonal subset. If $\boldsymbol{y}$ is a linear combination of the vectors in S, that is,*

$$\boldsymbol{y} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_p\boldsymbol{v}_p,$$

then

$$c_i = \frac{\boldsymbol{v}_i \cdot \boldsymbol{y}}{\boldsymbol{v}_i \cdot \boldsymbol{v}_i}, \quad \textit{called the } \textbf{\textit{Fourier coefficients}}.$$

Proof. Taking the dot product, we get

$$\boldsymbol{v}_i \cdot \boldsymbol{y} = \boldsymbol{v}_i \cdot (c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_p\boldsymbol{v}_p) = c_i(\boldsymbol{v}_i \cdot \boldsymbol{v}_i).$$

Since $\boldsymbol{v}_i \cdot \boldsymbol{v}_i \neq 0$, dividing both sides by $\boldsymbol{v}_i \cdot \boldsymbol{v}_i$ gives the formula. □

Example 4.5.2. Let $S = \{\boldsymbol{v}_1 = (1,2,3), \boldsymbol{v}_2 = (1,1,-1), \boldsymbol{v}_3 = (-5,4,-1)\}$, which was shown in Example 4.2.6 to be an orthogonal subset of $\mathbb{R}^3$. By Theorem 4.2.7, we see that S is, in fact, an orthogonal basis for $\mathbb{R}^3$. Let $\boldsymbol{y} = (-4,8,10)$. Since $\boldsymbol{y} \in \mathbb{R}^3$, $\boldsymbol{y}$ is a linear combination of the vectors of S. By Theorem 4.5.1, we have

$$\begin{aligned}\boldsymbol{y} &= \left(\frac{\boldsymbol{v}_1 \cdot \boldsymbol{y}}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right)\boldsymbol{v}_1 + \left(\frac{\boldsymbol{v}_2 \cdot \boldsymbol{y}}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\right)\boldsymbol{v}_2 + \left(\frac{\boldsymbol{v}_3 \cdot \boldsymbol{y}}{\boldsymbol{v}_3 \cdot \boldsymbol{v}_3}\right)\boldsymbol{v}_3 = \frac{42}{14}\boldsymbol{v}_1 + \frac{-6}{3}\boldsymbol{v}_2 + \frac{42}{42}\boldsymbol{v}_3 = 3\boldsymbol{v}_1 - 2\boldsymbol{v}_2 + \boldsymbol{v}_3 \\ &= 3\begin{bmatrix}1\\2\\3\end{bmatrix} - 2\begin{bmatrix}1\\1\\-1\end{bmatrix} + \begin{bmatrix}-5\\4\\-1\end{bmatrix} = \begin{bmatrix}-4\\8\\10\end{bmatrix}, \text{ that is, } [\boldsymbol{y}]_S = \begin{bmatrix}3\\-2\\1\end{bmatrix}.\end{aligned}$$

■

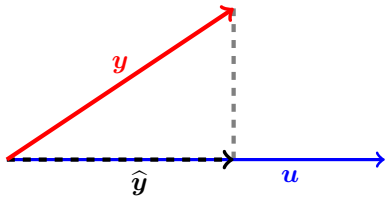
Definition 4.5.3. Let $\boldsymbol{y} \in F^n$ and $W \leq F^n$. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$ be an orthogonal basis for W. Then the **orthogonal projection** of $\boldsymbol{y}$ onto W is

$$\widehat{\boldsymbol{y}} = \text{proj}_W \boldsymbol{y} = \sum_{i=1}^{r} \frac{\boldsymbol{u}_i \cdot \boldsymbol{y}}{\boldsymbol{u}_i \cdot \boldsymbol{u}_i}\boldsymbol{u}_i = \frac{\boldsymbol{u}_1 \cdot \boldsymbol{y}}{\boldsymbol{u}_1 \cdot \boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{u}_2 \cdot \boldsymbol{y}}{\boldsymbol{u}_2 \cdot \boldsymbol{u}_2}\boldsymbol{u}_2 + \ldots + \frac{\boldsymbol{u}_r \cdot \boldsymbol{y}}{\boldsymbol{u}_r \cdot \boldsymbol{u}_r}\boldsymbol{u}_r.$$

When the subspace W is clear from context, $\text{proj}_W(\boldsymbol{y})$ is often abbreviated as $\widehat{\boldsymbol{y}}$.

Let $\boldsymbol{y}, \boldsymbol{u} \in F^n$ and $\boldsymbol{u} \neq \boldsymbol{0}$. Then the **orthogonal projection** of $\boldsymbol{y}$ onto $\boldsymbol{u}$ is $\text{proj}_{\boldsymbol{u}} \boldsymbol{y} = \text{proj}_W \boldsymbol{y}$, where $W = \text{Span}(\boldsymbol{u})$.

Intuitively, the orthogonal projection of $\boldsymbol{y}$ onto $\boldsymbol{u}$ is the shadow the arrow $\boldsymbol{y}$ casts on the line spanned along $\boldsymbol{u}$.



Example 4.5.4. Let $\boldsymbol{y} = \begin{bmatrix}1\\7\\3\end{bmatrix}$ and $\boldsymbol{u} = \begin{bmatrix}2\\1\\1\end{bmatrix}$. Then

$$\widehat{\boldsymbol{y}} = \text{proj}_{\boldsymbol{u}}(\boldsymbol{y}) = \frac{\boldsymbol{u}\cdot\boldsymbol{y}}{\boldsymbol{u}\cdot\boldsymbol{u}}\boldsymbol{u} = \frac{12}{6}\begin{bmatrix}2\\1\\1\end{bmatrix} = \begin{bmatrix}4\\2\\2\end{bmatrix}.$$

Note that

$$\boldsymbol{y}-\widehat{\boldsymbol{y}} = \begin{bmatrix}1\\7\\3\end{bmatrix} - \begin{bmatrix}4\\2\\2\end{bmatrix} = \begin{bmatrix}-3\\5\\1\end{bmatrix}, \qquad \boldsymbol{y} = \widehat{\boldsymbol{y}} + (\boldsymbol{y}-\widehat{\boldsymbol{y}}) = \begin{bmatrix}4\\2\\2\end{bmatrix} + \begin{bmatrix}-3\\5\\1\end{bmatrix},$$

and that

$$\widehat{\boldsymbol{y}}\cdot(\boldsymbol{y}-\widehat{\boldsymbol{y}}) = \begin{bmatrix}4\\2\\2\end{bmatrix}\cdot\begin{bmatrix}-3\\5\\1\end{bmatrix} = -12+10+2 = 0.$$

Therefore, $\boldsymbol{y}-\widehat{\boldsymbol{y}}$ and $\widehat{\boldsymbol{y}}$ are orthogonal, that is, $\boldsymbol{y}-\widehat{\boldsymbol{y}} \in W^{\perp}$. ∎

Theorem 4.5.5. *Let $\boldsymbol{y} \in F^n$ and let $W \leq F^n$. Then $\boldsymbol{y} - \text{proj}_W \boldsymbol{y} \in W^{\perp}$. In particular,*

$$\text{proj}_W \boldsymbol{y}\cdot(\boldsymbol{y}-\text{proj}_W \boldsymbol{y}) = 0.$$

Theorem 4.5.6 (The Orthogonal Decomposition Theorem)**.** *Let $\boldsymbol{y} \in F^n$ and $W \leq F^n$. Then there exists unique vectors $\boldsymbol{w}_1 \in W$ and $\boldsymbol{w}_2 \in W^{\perp}$ such that*

$$\boldsymbol{y} = \boldsymbol{w}_1 + \boldsymbol{w}_2.$$

Example 4.5.7. Let $\boldsymbol{u}_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}$, $\boldsymbol{u}_2 = \begin{bmatrix}-5\\4\\-1\end{bmatrix}$, and $\boldsymbol{y} = \begin{bmatrix}-9\\20\\-1\end{bmatrix}$. Let $W = \text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$. Since $\boldsymbol{u}_1\cdot\boldsymbol{u}_2 = 0$, $\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$ is an orthogonal basis for W. Using the orthogonal decomposition of $\boldsymbol{y}$ onto W, we can express $\boldsymbol{y}$ as a sum of a vector from W and a vector from $W^{\perp}$.

Now,

$$\boldsymbol{w}_1 = \widehat{\boldsymbol{y}} = \text{proj}_W \boldsymbol{y} = \frac{\boldsymbol{u}_1\cdot\boldsymbol{y}}{\boldsymbol{u}_1\cdot\boldsymbol{u}_1}\boldsymbol{u}_1 + \frac{\boldsymbol{u}_2\cdot\boldsymbol{y}}{\boldsymbol{u}_2\cdot\boldsymbol{u}_2}\boldsymbol{u}_2 = \frac{28}{14}\begin{bmatrix}1\\2\\3\end{bmatrix} + \frac{126}{42}\begin{bmatrix}-5\\4\\-1\end{bmatrix} = \begin{bmatrix}-13\\16\\3\end{bmatrix}.$$

Thus,

$$\boldsymbol{w}_2 = \boldsymbol{y}-\widehat{\boldsymbol{y}} = \begin{bmatrix}-9\\20\\-1\end{bmatrix} - \begin{bmatrix}-13\\16\\3\end{bmatrix} = \begin{bmatrix}4\\4\\-4\end{bmatrix}.$$

Note that $\boldsymbol{w}_2 \cdot \boldsymbol{u}_1 = \boldsymbol{w}_2 \cdot \boldsymbol{u}_2 = 0$, that is, $\boldsymbol{w}_2 \in W^\perp$. Therefore,

$$\boldsymbol{y} = \boldsymbol{w}_1 + \boldsymbol{w}_2 = \begin{bmatrix} -13 \\ 16 \\ 3 \end{bmatrix} + \begin{bmatrix} 4 \\ 4 \\ -4 \end{bmatrix} = \begin{bmatrix} -9 \\ 20 \\ -1 \end{bmatrix}. \qquad \blacksquare$$

Theorem 4.5.8 (The Best Approximation Theorem)**.** *Let $W \leq F^n$, let $\boldsymbol{y} \in F^n$, and let $\widehat{\boldsymbol{y}} = \operatorname{proj}_W \boldsymbol{y}$. Then $\widehat{\boldsymbol{y}}$ is the closest point in W to $\boldsymbol{y}$, that is,*

$$\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\| \leq \|\boldsymbol{y} - \boldsymbol{w}\|$$

for all $\boldsymbol{w} \in W$.

The vector $\widehat{\boldsymbol{y}}$ is called the **best approximation** to $\boldsymbol{y}$ in W.

Proof. Let $\boldsymbol{w} \in W$. Then

$$\boldsymbol{y} - \boldsymbol{w} = (\boldsymbol{y} - \widehat{\boldsymbol{y}}) + (\widehat{\boldsymbol{y}} - \boldsymbol{w}),$$

where $\boldsymbol{y} - \widehat{\boldsymbol{y}} \in W^\perp$ and $\widehat{\boldsymbol{y}} - \boldsymbol{w} \in W$, by the Orthogonal Decomposition Theorem. Then by the Pythagorean Theorem,

$$\|\boldsymbol{y} - \boldsymbol{w}\|^2 = \|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|^2 + \|\widehat{\boldsymbol{y}} - \boldsymbol{w}\|^2.$$

Therefore, $\|\boldsymbol{y} - \widehat{\boldsymbol{y}}\|^2 \leq \|\boldsymbol{y} - \boldsymbol{w}\|^2$. $\square$

Example 4.5.9. The distance from a point $\boldsymbol{y} \in \mathbb{R}^n$ to a subspace W is defined as the distance from $\boldsymbol{y}$ to the nearest point in W, the best approximation. Continuing with Example 4.5.7, the vector $\widehat{\boldsymbol{y}} = \begin{bmatrix} -13 \\ 16 \\ 3 \end{bmatrix}$ is the closest vector in W to the vector $\boldsymbol{y}$. Hence, $\operatorname{dist}(W, \boldsymbol{y}) = \|\boldsymbol{y} - \widehat{\boldsymbol{y}}\| = 4\sqrt{1^2 + 1^2 + (-1)^2} = \boxed{4\sqrt{3}}$. $\blacksquare$

Let H be a flat in F^n passing through the vector $\boldsymbol{h}$. Can we compute the distance between H and some vector $\boldsymbol{y}$? Modifying the Best Approximation Theorem, we can because we can translate H by the vector $-\boldsymbol{h}$ so that it becomes a subspace, say W. If we also translate $\boldsymbol{y}$ by $-\boldsymbol{h}$, then we get

$$\operatorname{dist}(H, \boldsymbol{y}) = \operatorname{dist}(W, \boldsymbol{y} - \boldsymbol{h}) = \|(\boldsymbol{y} - \boldsymbol{h}) - \widehat{(\boldsymbol{y} - \boldsymbol{h})}\| = \|(\boldsymbol{y} - \widehat{\boldsymbol{y}}) - (\boldsymbol{h} - \widehat{\boldsymbol{h}})\|,$$

since vector translation is an isometry and orthogonal projections are linear transformations.

It can be shown that $\operatorname{proj}_W : F^n \to F^n$ for some subspace $W \leq F^n$ is a linear transformation. Therefore, there must be some matrix that represent proj_W. Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_r$ denote an orthonormal basis for W. Let $Q = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \ldots & \boldsymbol{u}_r \end{bmatrix}$ be the $n \times r$ matrix whose columns vectors are this orthonormal basis. Then $\operatorname{proj}_W(\boldsymbol{y}) = QQ^\top(\boldsymbol{y})$, that is, $QQ^\top$ is the standard matrix representation for proj_W. Note that $QQ^\top$ can be thought of as a generalization of the outer product of vectors.

Exercises

(Go to Solutions)

For Exercises 1-5, use inner products to express the vector $\boldsymbol{y}$, listed first, as a linear combination of the orthogonal basis $\mathcal{B}$, listed second. Find the coordinate vector $[\boldsymbol{y}]_{\mathcal{B}}$.

♠ 1. $\begin{bmatrix} -2 \\ -6 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$

♠ 2. $\begin{bmatrix} -1 \\ 10 \\ 11 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

♠ 3. $\begin{bmatrix} 17 \\ -8 \\ 4 \end{bmatrix}, \left\{ \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \right\}$

♠ 4. $\begin{bmatrix} -6 \\ 9 \\ 8 \\ -5 \end{bmatrix}, \left\{ \begin{bmatrix} 3 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right\}$

♠ 5. $\begin{bmatrix} -19 \\ 18 \\ -12 \\ -15 \\ -13 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 6 \\ -6 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -2 \\ 0 \\ 5 \\ 5 \end{bmatrix} \right\}$

For Exercises 6-10, for the vector $\boldsymbol{y}$, listed first, and vector space $W = \text{Span}\{\mathcal{B}\}$, where $\mathcal{B}$ is an orthogonal basis, listed second, find $\text{proj}_W(\boldsymbol{y})$. Find the orthogonal decomposition of $\boldsymbol{y}$, that is, $\boldsymbol{y} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ where $\boldsymbol{w}_1 \in W$ and $\boldsymbol{w}_2 \in W^{\perp}$.

♠ 6. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$

♠ 7. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

♠ 8. $\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \left\{ \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \right\}$

♠ 9. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \left\{ \begin{bmatrix} 3 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right\}$

♠ 10. $\begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 6 \\ -6 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -2 \\ 0 \\ 5 \\ 5 \end{bmatrix} \right\}$

For Exercises 11-15, for the vector $\boldsymbol{y}$, listed first, and vector space $W = \text{Span}\{\mathcal{B}\}$, where $\mathcal{B}$ is an orthogonal basis, listed second, find the distance between $\boldsymbol{y}$ and W.

♠ 11. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$

♠ 12. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \right\}$

♠ 13. $\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \left\{ \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix} \right\}$

♠ 14. $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \left\{ \begin{bmatrix} 3 \\ 4 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \\ -5 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \right\}$

♠ 15. $\begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}, \left\{ \begin{bmatrix} 1 \\ 6 \\ -6 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 7 \\ -2 \\ 0 \\ 5 \\ 5 \end{bmatrix} \right\}$

16. Let V be an inner product space. Prove the *Parallelogram Law*:[†] For all $\boldsymbol{u},\ \boldsymbol{v} \in V$, $\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2 = 2\|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|^2$.

[†]This law gets its name from Euclidean space because the equality relates the lengths of the four sides of a parallelogram with its two diagonals.

"Most of the fundamental ideas of science are essentially simple, and may, as a rule, be expressed in a language comprehensible to everyone." – Albert Einstein

4.6 The Fundamental Theorem of Linear Algebra

Example 4.6.1.

Let $A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Thus, $\operatorname{Nul} A$ is a 3-dimensional subspace of $\mathbb{R}^5$ and $\operatorname{nullity} A = 3$. Likewise, $\operatorname{Col} A$ is a 2-dimensional subspace of $\mathbb{R}^3$ and $\operatorname{rank} A = 2$. ■

Example 4.6.2.

Let $A = \begin{bmatrix} 1 & 3 & 3 & 2 & -9 \\ -2 & -2 & 2 & -8 & 2 \\ 2 & 3 & 0 & 7 & 1 \\ 3 & 4 & -1 & 11 & -8 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -3 & 5 & 0 \\ 0 & 1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$. Thus, $\operatorname{Nul} A$ is a 2-dimensional subspace of $\mathbb{R}^5$ and $\operatorname{nullity} A = 2$. Also, $\operatorname{Col} A$ is a 3-dimensional subspace of $\mathbb{R}^4$ and $\operatorname{rank} A = 3$. ■

The rank of any matrix is the number of pivot positions. Suppose that A is an $m \times n$ matrix with p pivots. Then $\operatorname{rank} A = p$. This also implies that the number of non-pivot columns in A is $n - p$. Consider the homogeneous system $A\boldsymbol{x} = \boldsymbol{0}$. The number of non-pivot columns is the same as the number of free variables in the homogeneous system. As we have also seen, each free variable corresponds to a basis element for $\operatorname{Nul} A$. Thus, $\operatorname{nullity} A = n - p$. This proves the following important theorem.

Theorem 4.6.3 (The Rank-Nullity Theorem)**.** *If A is an $m \times n$ matrix, then*

$$\operatorname{rank} A + \operatorname{nullity} A = n.$$

Example 4.6.4. Let A be a 7×9 matrix. Suppose that $\operatorname{nullity} A = 2$. Then by the Rank-Nullity Theorem, $\operatorname{rank} A = 7$, that is, the mapping $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is surjective.

Is it possible for a 6×9 matrix A to have $\operatorname{nullity} A = 2$? No, because this implies that $\operatorname{rank} A = 7$. This is a contradiction because $\operatorname{rank} A \leq 6 = \dim \mathbb{R}^6$. In fact, the Rank-Nullity Theorem tells us that $\operatorname{nullity} A \geq 3$. ■

By the Rank-Nullity Theorem, we know that

$$\operatorname{rank} A + \operatorname{nullity} A = n.$$

Taking transposes, we all have

$$\dim(\operatorname{Row} A) + \dim(\operatorname{Nul}(A^\top)) = m.$$

Remark 4.6.5. The space $\operatorname{Nul}(A^\top)$ is often called the **left null space** of A because it consists of all $1 \times m$ row vectors $\boldsymbol{x}$ such that $\boldsymbol{x}A = \boldsymbol{0}$. ▼

Theorem 4.6.6. *For a matrix A, $\operatorname{rank} A = \operatorname{rank} A^\top$, that is, the dimension of the column and row space of A are the same.*

Proof. First, the $\operatorname{rank} A$ is equal to the number of pivot columns. Let B be the row reduced echelon form of A. Then $\dim(\operatorname{Row} A) = \dim(\operatorname{Row} B)$, where the second dimension is the same as the number of nonzero rows. But this equals the number of pivots. □

Theorem 3.3.11 (The Nonsingular Matrix Theorem–Continued) *Let A be a square $n \times n$ matrix. Then the following statements are equivalent:*

(i) *A is a nonsingular matrix.*

⋮

(xxi) $\operatorname{Col}(A) = F^n$

(xxii) $\operatorname{nullity}(A) = 0$

(xxiii) $\operatorname{Nul}(A) = \mathbf{0}$

(xxiv) $\operatorname{corank}(A) = n$

(xxv) $\operatorname{Row}(A) = F^n$

(xxvi) $\operatorname{conullity}(A) = 0$

(xxvii) $\operatorname{Lnl}(A) = \mathbf{0}$

We establish the four fundamental spaces of an $m \times n$ matrix A. Say A has p pivots. We have seen already three of these four fundamental spaces many times already:

(i) $\operatorname{Nul}(A) = \{\boldsymbol{x} \mid A\boldsymbol{x} = \mathbf{0}\} \leq F^n$, the **null space** of A (the set of column vectors which produce zero with A when multiplied on the right), $\operatorname{nullity}(T) = \operatorname{nullity}(A) = \dim \operatorname{Nul}(A) = n - p$.

(ii) $\operatorname{Col}(A) = \{A\boldsymbol{x} \mid \boldsymbol{x} \in F^n\} \leq F^m$, the **column space** of A (the span of the column vectors of A), $\operatorname{rank}(T) = \operatorname{rank}(A) = \dim \operatorname{Col}(A) = p$.

(iii) $\operatorname{Row}(A) = \{\boldsymbol{y}^\top \boldsymbol{A} \mid \boldsymbol{y} \in F^m\} = \operatorname{Col}(A^\top) \leq F^n$, the **row space** of A (the span of the rows vectors of A), $\operatorname{corank}(T) = \operatorname{corank}(A) = \dim \operatorname{Row}(A) = p$.

(iv) $\operatorname{Lnl}(A) = \{\boldsymbol{y} \mid \boldsymbol{y}^\top A = \mathbf{0}^\top\} = \operatorname{Nul}(A^\top) \leq F^m$, the **left null space** of A (the set of row vectors which produce zero with A when multiplied on the left), $\operatorname{conullity}(T) = \operatorname{conullity}(A) = \dim \operatorname{Lnl}(A) = m - p$.

The left null space essentially measures how much the column vectors of A do not span F^m, much in the same way the null space measures how much the column vectors are not linearly independent. We have seen that for a matrix A, the columns of A are linearly independent if and only if the null space of A is trivial. More generally, the dimension of the null space of A, its nullity, measures the size of the null space but also counts the number of free variables in the linear system $A\boldsymbol{x} = \boldsymbol{b}$. This is, of course, the number of non-pivot columns $n - p$. Similarly, the columns of A span F^m if and only if the left null space of A is trivial. More generally, the dimension of the left null space of A, its conullity, measures the size of the left null space but also counts the number of rows of zeros in the row-reduced echelon form U of A. This is, of course, the number of non-pivot rows $m - p$. The presence of a row of zeros in U allows the possibilities of inconsistent linear systems $A\boldsymbol{x} = \boldsymbol{b}$ and is dependent on the choice of $\boldsymbol{b}$. Essentially, the left null space is the subspace of F^m of those vectors $\boldsymbol{b}$ which "definitely" make $A\boldsymbol{x} = \boldsymbol{b}$ inconsistent.

Example 4.6.7. Let $A = \begin{bmatrix} 3 & -3 & -2 \\ -5 & 4 & 3 \\ 1 & -5 & -2 \end{bmatrix}$. Note that

$$\begin{bmatrix} 7 & 4 & -1 \end{bmatrix} \begin{bmatrix} 3 & -3 & -2 \\ -5 & 4 & 3 \\ 1 & -5 & -2 \end{bmatrix} = \begin{bmatrix} 21-20-1 & -21+16+5 & -14+12+2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} = \mathbf{0}^\top.$$

Therefore, $\boldsymbol{b} = \begin{bmatrix} 7 \\ 4 \\ -1 \end{bmatrix} \in \mathrm{Lnl}(A)$. Next, if we try to solve the linear system $A\boldsymbol{x} = \boldsymbol{b}$, we see that

$$\left[\begin{array}{ccc|c} 3 & -3 & -2 & 7 \\ -5 & 4 & 3 & 4 \\ 1 & -5 & -2 & -1 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -5 & -2 & -1 \\ 3 & -3 & -2 & 7 \\ -5 & 4 & 3 & 4 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -5 & -2 & -1 \\ 0 & 12 & 4 & 10 \\ 0 & -21 & -7 & -1 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & -5 & -2 & -1 \\ 0 & 3 & 1 & 5/2 \\ 0 & 3 & 1 & 1/7 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -5 & -2 & -1 \\ 0 & 3 & 1 & 5/2 \\ 0 & 0 & 0 & -33/14 \end{array}\right]$$

Therefore, $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, as was to be expected. ∎

For all the other fundamental spaces, we have an algorithm to find a basis for the subspace by row-reducing A. We provide one here. To find a basis for $\mathrm{Lnl}(A)$, augment A with the identity matrix I_m (of the appropriate size) and row reduce the matrix $[A \mid I_m]$. Let U be the row-reduced echelon form of A and let $E \in F^{m\times m}$ be a product of elementary matrices which transform A into U. Then $[A \mid I_m] \sim [U \mid E]$. Suppose that U has a row of zeros in ith row. Let $\boldsymbol{\varepsilon}_i$ be the ith row of E. Since $EA = U$, the ith row of U is factored as $\boldsymbol{\varepsilon}_i A$. As this is a row of zeros, $\boldsymbol{\varepsilon}_i \in \mathrm{Lnl}(A)$. As E is a nonsingular matrix, its set of row vectors is linearly independent, as is any subset of row vectors. Considering that $\dim \mathrm{Lnl}(A) = \mathrm{conullity}(A) = m - p =$ the number of zero rows, this independent set of row vectors must be a basis for $\mathrm{Lnl}(A)$.

Example 4.6.8. Continuing with Example 4.6.7,

$$[A \mid I_3] = \left[\begin{array}{ccc|ccc} 3 & -3 & -2 & 1 & 0 & 0 \\ -5 & 4 & 3 & 0 & 1 & 0 \\ 1 & -5 & -2 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -5 & -2 & 0 & 0 & 1 \\ 0 & 3 & 1 & 1/4 & 0 & -3/4 \\ 0 & 0 & 0 & -1/4 & -1/7 & 1/28 \end{array}\right]$$

Hence, the third row corresponds to a row of zeros. Thus, the row vector on the right side of the augmentation is a basis element of $\mathrm{Lnl}(A)$, namely, $\begin{bmatrix} -1/4 \\ -1/7 \\ 1/28 \end{bmatrix}$. We could replace this vector in the spanning set with any

nonzero scalar multiple, namely $\boldsymbol{b} = -28\begin{bmatrix} -1/4 \\ -1/7 \\ 1/28 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -1 \end{bmatrix}$. Thus, $\text{Lnl}(A) = \text{Span}\{(7, 4, -1)\}$. ∎

We present next the Fundamental Theorem of Linear Algebra, which is essentially a summary of many important results.

Theorem 4.6.9 (The Fundamental Theorem of Linear Algebra)**.** *Let A be an $m \times n$ matrix in $F^{m\times n}$. Then every vector $\boldsymbol{v} \in F^n$ can be decomposed uniquely as a sum of vectors:*

$$\boldsymbol{v} = \boldsymbol{v}_{\text{Row}(A)} + \boldsymbol{v}_{\text{Nul}(A)}$$

where $\boldsymbol{v}_{\text{Row}(A)} \in \text{Row}(A)$ and $\boldsymbol{v}_{\text{Nul}(A)} \in \text{Nul}(A)$. In particular, $\text{Row}(A)^\perp = \text{Nul}(A)$. Furthermore, if $A\boldsymbol{v} = \boldsymbol{b}$, then $\boldsymbol{v}_{\text{Row}(A)}$ is the particular solution to the linear system $A\boldsymbol{x} = \boldsymbol{b}$ of shortest length and $\boldsymbol{v}_{\text{Nul}(A)}$ is a solution to the associated homogeneous system $A\boldsymbol{x} = \boldsymbol{0}$.

Likewise, every vector $\boldsymbol{w} \in F^m$ can be decomposed uniquely as a sum of vectors:

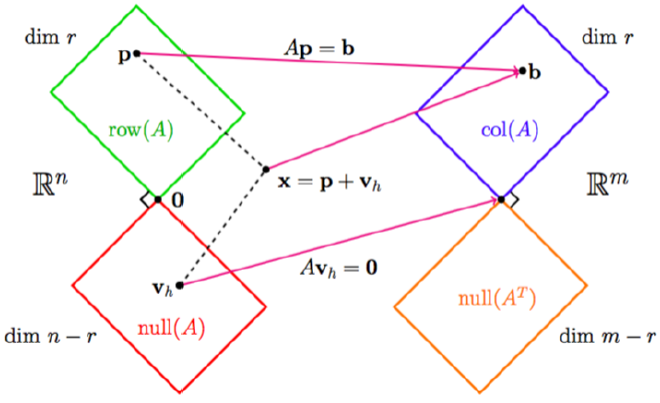
$$\boldsymbol{w} = \boldsymbol{w}_{\text{Col}(A)} + \boldsymbol{w}_{\text{Lnl}(A)}$$

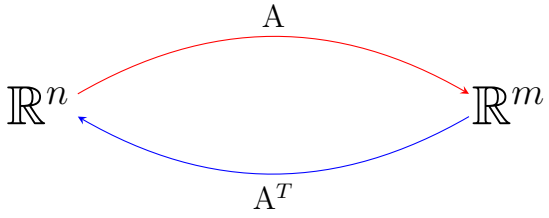
where $\boldsymbol{w}_{\text{Col}(A)} \in \text{Col}(A)$ and $\boldsymbol{w}_{\text{Lnl}(A)} \in \text{Lnl}(A)$. In particular, $\text{Col}(A)^\perp = \text{Lnl}(A)$. Furthermore, if $A^\top\boldsymbol{w} = \boldsymbol{b}$, then $\boldsymbol{w}_{\text{Col}(A)}$ is the particular solution to the linear system $A^\top\boldsymbol{y} = \boldsymbol{b}$ of shortest length and $\boldsymbol{w}_{\text{Lnl}(A)}$ is a solution to the associated homogeneous system $A^\top\boldsymbol{y} = \boldsymbol{0}$.

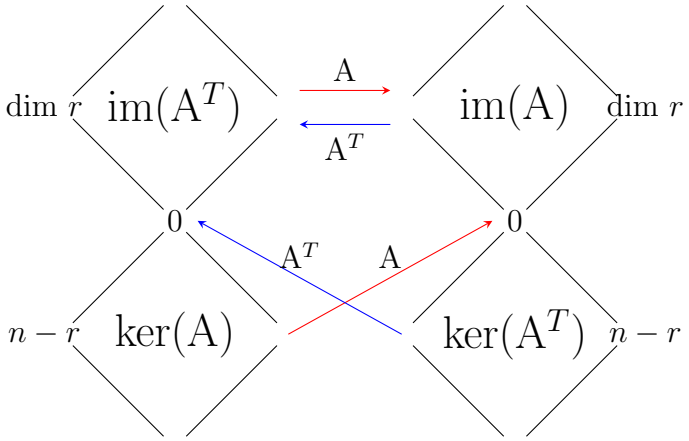
Finally,

$$\text{rank}(A) + \text{nullity}(A) = n, \qquad \text{corank}(A) + \text{conullity}(A) = m, \qquad \text{rank}(A) = \text{corank}(A).$$

This image needs to be replaced as it was ripped from another textbook. Also, the notion needs to be congruent with our presentation here.







Example 4.6.10. Solve the linear system $\begin{cases} 3x_1 - 3x_2 - 2x_3 = 3 \\ -5x_1 + 4x_2 + 3x_3 = -4 \\ x_1 - 5x_2 - 2x_3 = 5 \end{cases}$ by finding the particular solution contained in the row space of the coefficient matrix.

$$\left[\begin{array}{ccc|c} 3 & -3 & -2 & 3 \\ -5 & 4 & 3 & -4 \\ 1 & -5 & -2 & 5 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1/3 & 0 \\ 0 & 1 & 1/3 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus, the linear system is consistent and we see that

$$\boldsymbol{x} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix},$$

is the general solution to the linear system where $(1,-1,3) \in \mathrm{Nul}(A)$. But $(0,-1,0) \notin \mathrm{Row}(A)$. We know this since $(0,-1,0)\cdot(1,-1,3) = 1 \neq 0$. On the other hand, we may solve the equation $\boldsymbol{x}_{\mathrm{Row}(A)} \cdot (1,-1,3) = 0$.

$$\boldsymbol{x}_{\mathrm{Row}(A)} \cdot \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \left(\begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \right) \cdot \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = 0$$

$$(0+1+0) + t(1+1+9) = 0 \quad \Rightarrow \quad 1 + 11t = 0 \quad \Rightarrow \quad t = -\frac{1}{11}.$$

Hence, $\boldsymbol{x}_{\mathrm{Row}(A)} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \frac{1}{11} \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1/11 \\ -10/11 \\ -3/11 \end{bmatrix}$. It can be verified that $x_{\mathrm{Row}(A)}$ is both a solution to the linear system and a member of $\mathrm{Row}(A)$ (since it is orthogonal to $\mathrm{Nul}(A)$). ∎

Example 4.6.11. Solve the matrix equation $\begin{bmatrix} 2 & -6 & 0 & 0 & 0 \\ 3 & -7 & 4 & 14 & 4 \\ 4 & -19 & -14 & -49 & -14 \\ 3 & -6 & 6 & 21 & 6 \\ -1 & 2 & -2 & -7 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -72 \\ -78 \\ -249 \\ -63 \\ 21 \end{bmatrix}$ by finding the particular solution contained in the row space of the coefficient matrix.

We quickly solve the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$. We do this by row reducing the augmented matrix $[A \mid \boldsymbol{b}]$ to RREF. We then grab the particular solution associated to setting all free variables to zero and also grab the basis to the null space of the coefficient matrix from the RREF:

$$\left[\begin{array}{ccccc|c} 2 & -6 & 0 & 0 & 0 & -72 \\ 3 & -7 & 4 & 14 & 4 & -78 \\ 4 & -19 & -14 & -49 & -14 & -249 \\ 3 & -6 & 6 & 21 & 6 & -63 \\ -1 & 2 & -2 & -7 & -2 & 21 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 0 & 6 & 21 & 6 & 9 \\ 0 & 1 & 2 & 7 & 2 & 15 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Hence,

$$\mathrm{Nul}(A) = \mathrm{Span}\left\{\begin{bmatrix} -6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -21 \\ -7 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -6 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$$

and

$$\boldsymbol{x} = \begin{bmatrix} 9 \\ 15 \\ 0 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} -6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} -21 \\ -7 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -6 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 9-6r-21s-6t \\ 15-2r-7s-2t \\ r \\ s \\ t \end{bmatrix}.$$

To find the unique solution $\boldsymbol{x}_{\mathrm{Row}(A)} \in \mathrm{Row}(A)$, we utilize the fact that $\mathrm{Row}(A)^\top = \mathrm{Nul}(A)$. Hence, $\boldsymbol{x}_{\mathrm{Row}(A)}$ will be orthogonal to each vector in the basis of $\mathrm{Nul}(A)$:

$$\begin{cases} \boldsymbol{x}_{\mathrm{Row}(A)} \cdot (-6,-2,1,0,0) = 0 \\ \boldsymbol{x}_{\mathrm{Row}(A)} \cdot (-21,-7,0,1,0) = 0 \\ \boldsymbol{x}_{\mathrm{Row}(A)} \cdot (-6,-2,0,0,1) = 0 \end{cases} \sim \begin{cases} -6(9-6r-21s-6t) - 2(15-2r-7s-2t) + r = 0 \\ -21(9-6r-21s-6t) - 7(15-2r-7s-2t) + s = 0 \\ -6(9-6r-21s-6t) - 2(15-2r-7s-2t) + t = 0 \end{cases}$$

$$\sim \begin{cases} 41r + 140s + 40t = 84 \\ 140r + 491 + 140t = 294 \\ 40r + 140s + 41t = 84 \end{cases}$$

Solving this linear system, we get

$$\left[\begin{array}{ccc|c} 41 & 140 & 40 & 84 \\ 140 & 491 & 140 & 294 \\ 40 & 140 & 41 & 84 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{84}{571} \\ 0 & 1 & 0 & \frac{294}{571} \\ 0 & 0 & 1 & \frac{84}{571} \end{array}\right].$$

Therefore,

$$\boldsymbol{x}_{\mathrm{Row}(A)} = \begin{bmatrix} 9 \\ 15 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \frac{84}{571}\begin{bmatrix} -6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{294}{571}\begin{bmatrix} -21 \\ -7 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \frac{84}{571}\begin{bmatrix} -6 \\ -2 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{571}\begin{bmatrix} -2043 \\ 6171 \\ 84 \\ 294 \\ 84 \end{bmatrix}. \qquad \blacksquare$$

Exercises

(Go to Solutions)

For Exercises 1-4, consider the matrix A and the bases for two of its fundamental subspaces, e.g., $\text{Col}(A)$ and $\text{Nul}(A)$. How many rows and columns does A have, that is, what parameters $m \times n$ describe A?

1. $\text{Col}(A) = \text{Span}\left\{\begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \end{bmatrix}\right\}$, $\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

2. $\text{Col}(A) = \text{Span}\left\{\begin{bmatrix} 2 \\ 3 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ 4 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ -5 \\ 6 \\ 4 \end{bmatrix}\right\}$, $\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

3. $\text{Row}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}\right\}$, $\text{Lnl}(A) = \text{Span}\left\{\begin{bmatrix} 7 \\ 1 \\ -2 \end{bmatrix}\right\}$

4. $\text{Row}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right\}$, $\text{Lnl}(A) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1/4 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -3/4 \end{bmatrix}\right\}$

For Exercises 5-7, compute a basis for the left null space $\text{Lnl}(A)$ where A is the matrix provided. Answers may vary.

♠ 5. $\begin{bmatrix} -3 & 2 & 1 \\ 7 & -6 & -5 \\ -7 & 4 & 1 \end{bmatrix}$

♠ 6. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \\ 9 & 10 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 2 & -2 & -24 & -46 & 114 \\ 8 & -8 & -26 & -44 & 106 \\ -7 & 7 & 5 & 3 & -4 \\ -2 & 2 & 10 & 18 & -44 \end{bmatrix}$

For Exercises 8-10, solve the linear system $A\boldsymbol{x} = \boldsymbol{b}$ by finding a particular solution in the row space of the coefficient matrix, that is, find $\boldsymbol{x}_{\text{Row}(A)}$.

♠ 8. $\begin{bmatrix} 2 & 3 \\ -4 & -6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$

♠ 9. $\begin{bmatrix} 1 & 0 & -3 \\ -2 & 3 & 0 \\ -1 & -2 & 7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -8 \\ 4 \\ 16 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 6 & 6 & 1 & 5 \\ 1 & 1 & 0 & 1 \\ -2 & -2 & -4 & 2 \\ 0 & 0 & -3 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 9 \\ 2 \\ 8 \\ 9 \end{bmatrix}$

"Anyone who imagines that bliss is normal is going to waste a lot of time running around shouting that he has been robbed." – Jenkins Lloyd Jones

4.7 The Gram-Schmidt Algorithm

Using orthogonal projections, we can construct an orthogonal basis from a given basis. If desired, we can also construct an orthonormal basis by normalizing.

Theorem 4.7.1 (The Gram-Schmidt Process)**.** *Let $\mathcal{B} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p\}$ be a basis for a subspace $W \leq F^n$. Define recursively the vectors*

$$\begin{aligned}
\boldsymbol{v}_1 &= \boldsymbol{x}_1 \\
\boldsymbol{v}_2 &= \boldsymbol{x}_2 - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{x}_2}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 \\
\boldsymbol{v}_3 &= \boldsymbol{x}_3 - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{x}_3}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{x}_3}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 \\
&\vdots \\
\boldsymbol{v}_p &= \boldsymbol{x}_p - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{x}_p}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \ldots - \frac{\boldsymbol{v}_{p-1} \cdot \boldsymbol{x}_p}{\boldsymbol{v}_{p-1} \cdot \boldsymbol{v}_{p-1}}\boldsymbol{v}_{p-1}
\end{aligned}$$

Then $\mathcal{C} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\}$ is an orthogonal basis of W. In addition,

$$\text{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\} = \text{Span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \qquad \forall k.$$

Example 4.7.2. Let $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\boldsymbol{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$, which as a set is linearly independent. Let $W = \text{Span}\{\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3\}$. Then using the Gram-Schmidt procedure, we can construct an orthogonal basis of W.

$$\boldsymbol{v}_1 = \boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{v}_2 = \boldsymbol{x}_2 - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{x}_2}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$\boldsymbol{v}_3 = \boldsymbol{x}_3 - \frac{\boldsymbol{v}_1 \cdot \boldsymbol{x}_3}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\boldsymbol{v}_1 - \frac{\boldsymbol{v}_2 \cdot \boldsymbol{x}_3}{\boldsymbol{v}_2 \cdot \boldsymbol{v}_2}\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1/2}{12/16}\begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} - \begin{bmatrix} -3/6 \\ 1/6 \\ 1/6 \\ 1/6 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

Note that

$$\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = -3/4+1/4+1/4+1/4 = 0, \quad \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = 0-2/3+1/3+1/3 = 0, \quad \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = 0-2/3+1/3+1/3 = 0.$$

Therefore, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is an orthogonal basis for W. If we normalize:

$$\boldsymbol{u}_1 = \frac{1}{\sqrt{4}} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}, \quad \boldsymbol{u}_2 = \sqrt{\frac{4}{3}} \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix} = \begin{bmatrix} -3/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \\ 1/\sqrt{12} \end{bmatrix}, \quad \boldsymbol{u}_3 = \sqrt{\frac{9}{6}} \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix},$$

then $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthonormal basis of W. ∎

Example 4.7.3. Let $\boldsymbol{u} = \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix}$ and $\boldsymbol{v} = \begin{bmatrix} 1 \\ 0 \\ -i \end{bmatrix}$. Let $W = \text{Span}\{\boldsymbol{u},\ \boldsymbol{v}\} \subseteq \mathbb{C}^n$. Compute an orthogonal basis for W.

The set $\{\boldsymbol{u},\ \boldsymbol{v}\}$ is a basis for W since the set is linearly independent ($\boldsymbol{v} \neq z\boldsymbol{u}$ for any scalar $z \in \mathbb{C}$). Applying the Gram-Schmidt process, we set $\boldsymbol{x}_1 = \boldsymbol{u}$ and

$$\boldsymbol{x}_2 = \boldsymbol{v} - \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}} \boldsymbol{u} = \begin{bmatrix} 1 \\ 0 \\ -i \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ i \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -i/2 \\ -i \end{bmatrix}.$$

Note that

$$\boldsymbol{x}_1 \cdot \boldsymbol{x}_2 = 1\left(\frac{1}{2}\right) - i\left(-\frac{i}{2}\right) + 0(-i) = \frac{1}{2} - \frac{1}{2} = 0.$$

Thus, $\{\boldsymbol{x}_1,\ \boldsymbol{x}_2\}$ is an orthogonal basis for W. ∎

Theorem 4.7.4 (The QR Factorization)**.** *If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix with orthonormal columns and* $\text{Col}\, A = \text{Col}\, Q$ *and R is an $n \times n$ upper triangular matrix with positive diagonal entries. In particular, R is necessarily invertible.*

The matrix Q is constructed by using an orthonormal basis provided by the Gram-Schmidt process, although some columns might need to be scaled by -1 to guarantee that the diagonal entries of R are positive. If $A = QR$, then it can be shown that $R = \begin{bmatrix} \langle \boldsymbol{q}_i,\ \boldsymbol{a}_j \rangle \end{bmatrix}$, where $\boldsymbol{q}_i$ and $\boldsymbol{a}_j$ denote the columns of Q and A, respectfully. With respect to the dot product, this gives that $Q^\top A = R$.

Example 4.7.5. Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. By Example 4.7.2, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ is an orthonormal basis for Col A where

$$Q = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{bmatrix} = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}.$$

Let

$$R = Q^\top A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 1/\sqrt{3} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}.$$

Therefore, $A = QR$. ∎

Exercises

(Go to Solutions)

For Exercises 1-7, apply the Gram-Schmidt algorithm to find an orthonormal basis for each subspace W provided.

♠ 1. Span $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$

♠ 2. Span $\left\{ \begin{bmatrix} 4 \\ -4 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -5 \end{bmatrix} \right\}$

♠ 3. Span $\left\{ \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$

♠ 4. Span $\left\{ \begin{bmatrix} 1 \\ 0 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 4 \\ 2 \end{bmatrix} \right\}$

♠ 5. Span $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix} \right\}$

♠ 6. Span $\left\{ \begin{bmatrix} 1 \\ -i \\ 2i \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1-i \end{bmatrix} \right\}$

♠ 7. Span $\left\{ \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1-i \\ 0 \\ 3+2i \\ i \end{bmatrix} \right\}$

For Exercises 8-12, compute the QR-factorizations of the matrix. Note that many of the column spaces of these matrices coincide with subspaces considered in Exercises 1-7.

♠ 8. $\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix}$

♠ 9. $\begin{bmatrix} 0 & -1 \\ 1 & 3 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 4 & 2 \\ -4 & 4 \\ 2 & -5 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \\ 1 & 2 \end{bmatrix}$

♠ 12. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ 3 & -3 & 3 \end{bmatrix}$

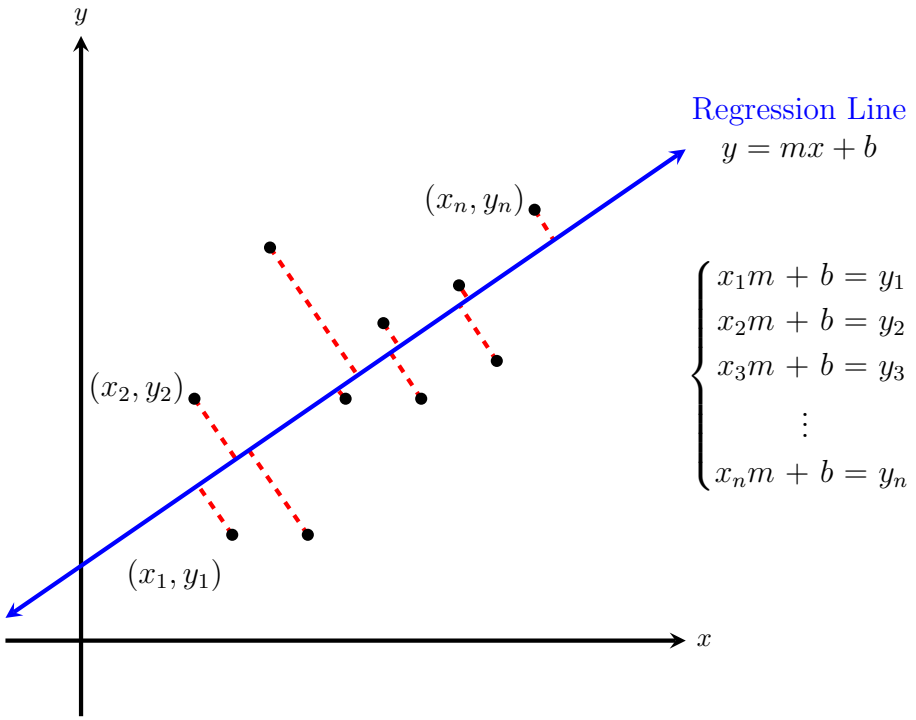
"Hypotheses are what we lack the least." – Henri Poincare

4.8 The Least Squares Problem

In the past, we have considered the matrix equation $A\boldsymbol{x} = \boldsymbol{b}$. We have learned how to determine if the equation has a solution and how to calculate the solution set when it is consistent. In fact, if $\boldsymbol{x}$ is a solution, then we can decompose

$$\boldsymbol{x} = \boldsymbol{x}_{\text{Row}(A)} + \boldsymbol{x}_{\text{Nul}(A)}$$

orthogonally. But what about when $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent? Well, the solution set is empty then, but what if we really needed it to have a solution? No manner of begging will change the fact that the matrix equation does not have a solution. For example, what if we want to find a line that is incident to all of the below points. As these points are not collinear, no such line exists. But is there a line that best fits the points, a so-called *regression line*?



Instead of asking, "Is there a vector $\boldsymbol{x}$ such that $A\boldsymbol{x} = \boldsymbol{b}$," we ask, "Is there a vector $\boldsymbol{x}$ such that $A\boldsymbol{x} \approx \boldsymbol{b}$?" By approximation here, we mean to find a vector $\boldsymbol{x}$ such that $\|A\boldsymbol{x} - \boldsymbol{b}\|$ is sufficiently small. This leads to the **least-squares problem**.

Definition 4.8.1. If A is an $m \times n$ matrix and $\boldsymbol{b} \in F^m$, a **least-squares solution** of $A\boldsymbol{x} = \boldsymbol{b}$ is a vector $\widehat{\boldsymbol{x}} \in F^n$ such that

$$\|\boldsymbol{b} - A\widehat{\boldsymbol{x}}\| \leq \|\boldsymbol{b} - A\boldsymbol{x}\|$$

for all $\boldsymbol{x} \in F^n$

Note that if $A\boldsymbol{x} = \boldsymbol{b}$ is consistent, then there exists a vector $\widehat{\boldsymbol{x}}$ such that $A\widehat{\boldsymbol{x}} = \boldsymbol{b}$ and $\|\boldsymbol{b} - A\widehat{\boldsymbol{x}}\| = \|\boldsymbol{0}\| = 0$. Thus, any solution is a least-square solution. On the other hand, if $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, then the matrix equation has no solution but must certainly still have a least-square solution. To find this least-squares solution, let us change our focus. Why is the linear system inconsistent and who should be blamed for it? Well, the vector $\boldsymbol{b}$, sort of. We have also seen the $\boldsymbol{b}$ can be orthogonally decomposed as

$$\boldsymbol{b} = \boldsymbol{b}_{\text{Col}(A)} + \boldsymbol{b}_{\text{Lnl}(A)}.$$

If $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, then it is because the left null space component of $\boldsymbol{b}$, that is, $\boldsymbol{b}_{\text{Lnl}(A)}$, is nontrivial. As we mentioned before, while the null space measures the ease of solving a consistent system (because it

increases the dimension of the solution set), the left null space measures the difficulty of the system being consistent (because more rows of zeros in the echelon form of A place more limitation on the choice of consistent $\boldsymbol{b}$).

Let A be an $m \times n$ matrix and $\boldsymbol{b} \in F^m$. Let $\boldsymbol{x} \in F^n$ be a generic vector (treat $\boldsymbol{x}$ as a variable vector). Then $A\boldsymbol{x}$ is a generic vector of the subspace $\operatorname{Col} A$. Let $\widehat{\boldsymbol{b}} = \operatorname{proj}_{\operatorname{Col} A} \boldsymbol{b} = \boldsymbol{b}_{\operatorname{Col}(A)}$. Then

$$\|\boldsymbol{b} - \widehat{\boldsymbol{b}}\| \leq \|\boldsymbol{b} - A\boldsymbol{x}\|$$

by the Best Approximation Theorem. Since $\widehat{\boldsymbol{b}} \in \operatorname{Col} A$, there exists some vector $\widehat{\boldsymbol{x}} \in F^n$ such that $A\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{b}}$. Therefore, every linear system has a least-squares solution $\widehat{\boldsymbol{x}}$, which is a solution to the linear system $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$, since for all $\boldsymbol{x}$ we have

$$\|\boldsymbol{b} - A\widehat{\boldsymbol{x}}\| \leq \|\boldsymbol{b} - A\boldsymbol{x}\|.$$

Of course, we also have that the linear system $A\boldsymbol{x} = \widehat{\boldsymbol{b}}$ is ALWAYS consistent.

If $\widehat{\boldsymbol{x}}$ is a least-squares solution of $A\boldsymbol{x} = \boldsymbol{b}$, then $\boldsymbol{b} - A\widehat{\boldsymbol{x}} = \boldsymbol{b} - \widehat{\boldsymbol{b}} \in (\operatorname{Col} A)^\perp = \operatorname{Lnl}(A)$. Hence, $\boldsymbol{b} - A\widehat{x} = \boldsymbol{b}_{\operatorname{Lnl}(A)}$. Thus, $A^\top(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \boldsymbol{0}$. Therefore, $A^\top A\widehat{\boldsymbol{x}} = A^\top \boldsymbol{b}$. In particular, every least-squares solution to $A\boldsymbol{x} = \boldsymbol{b}$ is a solution to the system of **normal equations**

$$A^\top A\widehat{\boldsymbol{x}} = A^\top \boldsymbol{b}. \tag{4.8.1}$$

Essentially, multiplying by $A^\top$ kills off the left null space component of $\boldsymbol{b}$, leaving the column space component behind. For complex linear systems, replace $^\top$ with *, per the usual.

Theorem 4.8.2. *The set of least-squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$ coincides with the solution set of the normal equations $A^\top A\widehat{\boldsymbol{x}} = A^\top \boldsymbol{b}$.*

Example 4.8.3. Find a least-squares solution of the inconsistent system $A\boldsymbol{x} = \boldsymbol{b}$ with

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}.$$

We begin by constructing the normal equations:

$$A^\top A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}; \qquad A^\top \boldsymbol{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Thus, $A^\top A\widehat{\boldsymbol{x}} = A^\top \boldsymbol{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}.$$

Since $A^\top A$ is invertible, we have

$$\widehat{\boldsymbol{x}} = (A^\top A)^{-1} A^\top \boldsymbol{b} = \frac{1}{84} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}. \qquad \blacksquare$$

Example 4.8.4. Find a least-squares solution for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \text{ and } \boldsymbol{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}.$$

$$A^\top A = \begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}; \qquad A^\top \boldsymbol{b} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}; \qquad \left[\begin{array}{cccc|c} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Therefore,

$$\widehat{\boldsymbol{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \qquad \blacksquare$$

As illustrated in the last example, a least-squares solution need not be unique.

Theorem 4.8.5. *Let A be an $m \times n$ matrix. Then the following are equivalent:*

(i) The equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least-squares solution for each $\boldsymbol{b} \in F^m$.

(ii) The columns of A are linearly independent.

(iii) The matrix $A^\top A$ is invertible.

In this situation, the unique least-squares solution has the form

$$\widehat{\boldsymbol{x}} = (A^\top A)^{-1} A^\top \boldsymbol{b}.$$

Example 4.8.6. The matrix A given in Example 4.8.3 has a unique least-squares solution for all $\boldsymbol{b} \in \mathbb{R}^m$. $\blacksquare$

Least-squares solutions also allow us to compute orthogonal projections without an orthogonal basis. For example, if $W = \text{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_r\} \leq F^n$, $A = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \ldots & \boldsymbol{v}_r \end{bmatrix}$, and $\boldsymbol{y} \in F^n$, then $W = \text{Col}(A)$ and $\widehat{\boldsymbol{y}} = \text{proj}_W(\boldsymbol{y}) = \text{proj}_{\text{Col}(A)}(\boldsymbol{y}) = \boldsymbol{y}_{\text{Col}(A)} = A\widehat{\boldsymbol{x}}$ where $\boldsymbol{x}$ is a least-squares solutions to $A\boldsymbol{x} = \boldsymbol{y}$.

Example 4.8.7. Let $W = \text{Col}(A)$, where A is defined in Example 4.8.3. Then $\widehat{\boldsymbol{b}} = \text{proj}_W(\boldsymbol{b}) = A\widehat{\boldsymbol{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$. Thus, we may find orthogonal projections by solving least-squares problem. $\blacksquare$

In the vein of the previous example, the standard matrix of the orthogonal projection from F^n onto the subspace W is $A(A^\top A)^{-1}A^\top$, where the columns of A correspond to a basis for W, since $\boldsymbol{b} \mapsto A(A^\top A)^{-1}A^\top \boldsymbol{b} = A\widehat{\boldsymbol{x}} = \widehat{b} = \text{proj}_W(\boldsymbol{b})$. Of course, if $A = QR$ is a QR factorization, then

$$A(A^\top A)^{-1}A^\top = QR((QR)^\top QR)^{-1}(QR)^\top = QR(R^\top Q^\top QR)^{-1}R^\top Q^\top = QR(R^\top R)^{-1}R^\top Q^\top = QR(R^{-1}(R^\top)^{-1})R^\top Q^\top = Q$$

This second form appears much simpler (and agrees with the formula we found in Section 4.5), but this is because know $A = QR$ means we have an orthonormal basis for W, namely the columns of Q. The advantage of $A(A\top A)^{-1}A^\top$ is that no orthogonalization is necessarily to determine proj_W.

Example 4.8.8. Let $A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}$ and $W = \text{Col}(A)$. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the orthogonal projection onto W. Then

$$[T] = A(A^\top A)^{-1}A^\top = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix} \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}$$

$$= \frac{1}{84} \begin{bmatrix} 20 & -4 \\ -2 & 34 \\ 4 & 16 \end{bmatrix} \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix} = \frac{1}{84} \begin{bmatrix} 80 & -8 & 16 \\ -8 & 68 & 32 \\ 16 & 32 & 20 \end{bmatrix} = \frac{1}{21} \begin{bmatrix} 20 & -2 & 4 \\ -2 & 17 & 8 \\ 4 & 8 & 5 \end{bmatrix}$$

Note that

$$\frac{1}{21} \begin{bmatrix} 20 & -2 & 4 \\ -2 & 17 & 8 \\ 4 & 8 & 5 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \frac{1}{21} \begin{bmatrix} 40+0+44 \\ -4+0+88 \\ 8+0+55 \end{bmatrix} = \frac{1}{21} \begin{bmatrix} 84 \\ 84 \\ 63 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}. \qquad \blacksquare$$

Theorem 4.8.9. *Let A be an $m \times n$ matrix with linearly independent columns. Let $A = QR$ be a QR factorization. Let $\boldsymbol{b} \in F^m$. Then the unique least-squares solution $\widehat{\boldsymbol{x}} \in F^n$ has the form*

$$\widehat{\boldsymbol{x}} = R^{-1}Q^\top \boldsymbol{b}.$$

Proof. Let $\widehat{\boldsymbol{x}} = R^{-1}Q^\top \boldsymbol{b}$. Then

$$A\widehat{\boldsymbol{x}} = A(R^{-1}Q^\top \boldsymbol{b}) = QR(R^{-1}Q^\top \boldsymbol{b}) = QQ^\top \boldsymbol{b} = \text{proj}_{\text{Col}\, Q}\, \boldsymbol{b},$$

since $QQ^\top$ is the standard matrix representation for $\text{proj}_{\text{Col}\, Q}$. Since $\text{Col}\, Q = \text{Col}\, A$, $\widehat{\boldsymbol{x}}$ is a least-squares solution of $A\boldsymbol{x} = \boldsymbol{b}$. By Theorem 4.8.5, this is the unique least-squares solution. $\square$

Exercises

(Go to Solutions)

For Exercises 1-3, compute the least square solutions for linear system.

♠ 1. $\left[\begin{array}{cc|c} 1 & -1 & 2 \\ 2 & 3 & -1 \\ 4 & 5 & 5 \end{array}\right]$

♠ 2. $\left[\begin{array}{cc|c} 0 & 1 & 3 \\ 1 & 0 & -2 \\ 2 & -1 & -7 \\ 1 & 0 & -2 \end{array}\right]$

♠ 3. $\left[\begin{array}{ccc|c} 2 & 0 & -1 & 0 \\ 1 & -2 & 2 & 6 \\ 2 & -1 & 0 & 0 \\ 0 & 1 & -1 & 6 \end{array}\right]$

For Exercises 4-5, compute the orthogonal projection $\text{proj}_W(\boldsymbol{b})$ using the method of least squares as in Example 4.8.7 where $W = \text{Col}(A)$, where A is the matrix provided first and $\boldsymbol{b}$ is the vector provided second.

♠ 4. $\left[\begin{array}{cc} 1 & -1 \\ 3 & 2 \\ -2 & 4 \end{array}\right], \left[\begin{array}{c} 4 \\ 1 \\ 3 \end{array}\right]$

♠ 5. $\left[\begin{array}{cc} 5 & 1 \\ 1 & 3 \\ 4 & -2 \end{array}\right], \left[\begin{array}{c} -4 \\ 2 \\ 3 \end{array}\right]$

For Exercises 6-7, compute the orthogonal projection $\text{proj}_W(\boldsymbol{b})$ using the method of least-squares. where a basis for W is provided first and the vector $\boldsymbol{b}$ is the vector provided second.

♠ 6. $\{(-1, 2, 1),\ (2, 2, 4)\},\ (1, -6, 1)$

♠ 7. $\{(2, 1, 1, 1),\ (1, 0, 1, 1),\ (-2, -1, 0, -1)\},\ (6, 3, 9, 6)$

For Exercises 8-9, find a basis for W and the standard matrix for the orthogonal projection onto W. Answers may vary.

♠ 8. W is the plane given by the equation $5x - 3y + z = 0$ in $\mathbb{R}^3$.

♠ 9. W is the line with parametric equations $x = 2t$, $y = -t$, and $z = 4t$ in $\mathbb{R}^3$.

10. If A is a nonsingular matrix, show that the unique least squares solution $\widehat{\boldsymbol{x}}$ to the linear system $A\boldsymbol{x} = \boldsymbol{b}$ is likewise the unique linear solution to $A\boldsymbol{x} = \boldsymbol{b}$.

11. Suppose that $A = QR$ is a QR factorization of A. If $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least-squares solution, show that $\widehat{x} = R^{-1}Q^{\top}\boldsymbol{b}$ is this solution.

Chapter 5

Determinants

"Our lives will depend upon the decisions which we make–for decisions determine destiny."
– Thomas S. Monson

5.1 Introduction to Determinants

In a previous section, we discussed the inverse of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$:

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The quantity $ad - bc$ was called the **determinant** of A, denoted $\det A$ or $|A|$. In this chapter we work to generalize this idea for any square matrix. Our strategy will be to define determinants recursively. We will remind the reader that, unlike the previous chapter, the theory of determinants is applicable to any field. Hence, F will denote an arbitrary field.

Definition 5.1.1. Let A be an $n \times n$ matrix. Define the (i, j)-**minor matrix** A_{ij} to be the $(n-1) \times (n-1)$ matrix which results by removing the ith row and jth column from A.

Example 5.1.2. Let $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$. Compute the $(1,1)-$, $(1,2)-$, $(2,2)-$, and $(3,3)$-minor matrix.

$$A_{11} = \begin{bmatrix} 4 & -1 \\ -2 & 0 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}, \quad A_{22} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad A_{33} = \begin{bmatrix} 1 & 5 \\ 2 & 4 \end{bmatrix}. \qquad \blacksquare$$

Definition 5.1.3. Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ be an $n \times n$ matrix. For $n = 1$, $A = \begin{bmatrix} r \end{bmatrix}$ for some $r \in F$. In this case, we define $\det A = r$. For $n > 1$, let

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + \ldots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{ij} \det A_{ij} \end{aligned}$$

The quantity $\det A$ is called the **determinant** of A.

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $\det A = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \det A_{11} - b \det A_{12} = a \begin{vmatrix} d \end{vmatrix} - b \begin{vmatrix} c \end{vmatrix} = ad - bc$, which agrees with our 2×2 determinant from before.

Example 5.1.4. Compute $\det A$ for $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

$$\begin{aligned}\det A &= \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix} = 1\begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 5\begin{vmatrix} 2 & -1 \\ 0 & 0 \end{vmatrix} + 0\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} \\ &= 1(4\cdot 0 - (-1)\cdot(-2)) - 5(2\cdot 0 - (-1)\cdot 0) + 0 = (0-2) - 5(0) = -2. \quad \blacksquare\end{aligned}$$

Definition 5.1.5. Let A be an $n \times n$ matrix. Define the (i,j)-**cofactor** C_{ij} as

$$C_{ij} = (-1)^{i+j} \det A_{ij}.$$

With this notation,

$$\det A = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}.$$

The following diagram may help you remember the plus or minus sign in the cofactors:

$$\begin{bmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \\ \vdots & & & \ddots \end{bmatrix}$$

Theorem 5.1.6 (Laplace Expansion)**.** *The determinant of an $n \times n$ matrix $A = \begin{bmatrix} a_{ij} \end{bmatrix}$ can be computed by a cofactor expansion across any row or down any column. The cofactor expansion across the ith row is*

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in} = \sum_{j=1}^{n} a_{ij}C_{ij}.$$

The cofactor expansion down the jth column is

$$\det A = a_{ij}C_{1j} + a_{2j}C_{2j} + \ldots + a_{nj}C_{nj} = \sum_{i=1}^{n} a_{ij}C_{ij}.$$

Example 5.1.7. Use the cofactor expansion across the third row to compute $\det A$, where $A = \begin{bmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{bmatrix}$.

$$\begin{aligned}\det A &= \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix} = 0\begin{vmatrix} 5 & 0 \\ 4 & -1 \end{vmatrix} - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix} \\ &= 2\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = 2(-1-0) = \boxed{-2}.\end{aligned}$$

$\blacksquare$

Example 5.1.8. Compute $\det A$, where $A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}$.

We will expand across the leftmost column to maximize the number of zero coefficients.

$$\begin{vmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{vmatrix} = 3 \begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} = 6 \begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}$$

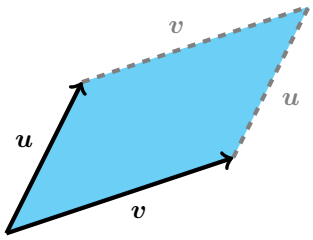
$$= 6 \left(\begin{vmatrix} 4 & -1 \\ -2 & 0 \end{vmatrix} - 2 \begin{vmatrix} 5 & 0 \\ -2 & 0 \end{vmatrix} \right) = 6(-2+0) = \boxed{-12}$$

■

The matrix in the last example is nearly triangular. The method in that example is easily adapted to prove the following.

Theorem 5.1.9. *If A is a triangular $n \times n$ matrix, then $\det A$ is the product of the entries on the main diagonal.*

Theorem 5.1.10. *If A is a 2×2 matrix over $\mathbb{R}$ or $\mathbb{C}$, the area of the parallelogram determined by the columns of A is $|\det A|$. If A is 3×3, the volume of the parallelepiped determined by the columns of A is $|\det A|$. The higher dimensional analogues also hold.*



Example 5.1.11. Calculate the area of the parallelogram determined by the points $(-2,-2)$, $(0,3)$, $(4,-1)$, and $(6,4)$.

We begin by translating the parallelogram such that $(-2,-2)$ is moved to the origin. This translation does not affect the area. We now consider the parallelogram with vertices $(0,0)$, $(2,5)$, $(6,1)$, and $(8,6)$. Note that this parallelogram is the one spanned by $(2,5)$ and $(6,1)$. Therefore,

$$\text{Area} = \text{abs} \begin{vmatrix} 2 & 6 \\ 5 & 1 \end{vmatrix} = |2 - 30| = \boxed{28}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-5, for the matrix $A = \begin{bmatrix} 5 & 7 & 12 \\ 0 & 3 & 4 \\ 9 & -1 & -6 \end{bmatrix}$, find the given minor.

1. A_{22} 2. A_{33} 3. A_{11} 4. A_{23} 5. A_{12}

For Exercises 6-10, for the matrix $B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -2 & -1 & 3 \\ 1 & 1 & 3 & 6 \\ 7 & -4 & 8 & 0 \end{bmatrix}$, find the given minor.

♠ 6. B_{11} ♠ 7. B_{13} ♠ 8. B_{31} ♠ 9. B_{33} ♠ 10. B_{34}

For Exercises 11-13, for the matrix $C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 3 & -1 \\ -2 & 4 & 0 & -4 \\ -1 & 3 & 2 & 5 \end{bmatrix}$, find the given minor.

11. C_{22} 12. C_{23} 13. C_{42}

For Exercises 14-16, for the matrix $D = \begin{bmatrix} i & 3i-2 & 2+5i & 7-i \\ 3+i & 4 & 7i-5 & 9-3i \\ 2+3i & 5i-1 & 1+9i & 4+5i \\ 7+2i & i-3 & 2+4i & 1+2i \end{bmatrix}$, find the given minor.

14. D_{24} 15. D_{33} 16. D_{14}

For Exercises 17-25, compute the determinant of the matrix.

♠ 17. $\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$ ♠ 18. $\begin{vmatrix} i & 1+i \\ 2-i & 3-4i \end{vmatrix}$ ♠ 19. $\begin{vmatrix} 3 & 2 \\ 1 & 5 \end{vmatrix} \pmod 7$

♠ 20. $\begin{vmatrix} 1 & 2 & -1 \\ -1 & 0 & 2 \\ 3 & 5 & 1 \end{vmatrix}$ ♠ 21. $\begin{vmatrix} 1 & 2 & 3 \\ 3 & -2 & 5 \\ 0 & 0 & 2 \end{vmatrix}$ ♠ 22. $\begin{vmatrix} 2 & 0 & 0 \\ 3 & 1 & 0 \\ 4 & 3 & 4 \end{vmatrix} \pmod 5$

♠ 23. $\begin{vmatrix} 1 & 1 & 2 & 4 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 2 & 4 \\ 1 & 2 & 3 & 4 \end{vmatrix} \pmod 5$

♠ 24. $\begin{vmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{vmatrix} \pmod 2$

25. $\begin{vmatrix} 5 & 4 & 4 & 2 \\ 2 & 4 & 6 & 8 \\ 0 & 1 & 2 & 3 \\ 3 & 3 & 3 & 9 \end{vmatrix}$

26. Compute $\det(A)(\boldsymbol{u} \cdot \boldsymbol{v})$ if:

$$A = \begin{bmatrix} \pi^2 & e \\ 5 & \sqrt{\pi} \end{bmatrix}, \quad \boldsymbol{u} = \begin{bmatrix} 5e^3 \\ 1 \\ \pi \end{bmatrix}, \quad \boldsymbol{v} = \begin{bmatrix} e^{-2} \\ 0 \\ \sqrt{\pi^3} \end{bmatrix}.$$

For Exercises 27-29, compute the area, volume, or hyper-volume of the parallelogram, parallelepiped, or hyper-parallelepiped spanned by the given set of vectors.

♠ 27. $\left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$

♠ 28. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 3 \\ 7 \end{bmatrix} \right\}$

♠ 29. $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 4 \\ 5 \end{bmatrix} \right\}$

"In most encounters we can determine the kind of experience we are going to have by how we respond."
– Wayne S. Peterson

5.2 Properties of Determinants

Let V and W be vector spaces over F. We have seen previously that linear maps $T : V \to W$ are those that preserve the linear operations, that is, $T(\boldsymbol{x} + \boldsymbol{y}) = T(\boldsymbol{x}) + T(\boldsymbol{y})$ and $T(c\boldsymbol{x}) = cT(\boldsymbol{x})$ for all $\boldsymbol{x}, \boldsymbol{y} \in V$ and $c \in F$. Let $V^n = \{(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n) \mid \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V\}$, that is, V^n is the set of all list of n vectors from V. This is itself a vector space with $\dim V^n = n \dim V$. When $V = F^m$, then $V^n \cong F^{m\times n}$, the space of all $m \times n$ matrices as each list in V^n can be identified with the columns vectors of an $m \times n$ matrix.

Definition 5.2.1. Let V and W be vector spaces over F. Let $n \in \mathbb{N}$. Let $B : V^n \to W$ be a function. We say that B is **multilinear** if for each i and choice of vector $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ the function

$$\boldsymbol{x} \mapsto f(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{i-1}, \boldsymbol{x}, \boldsymbol{v}_{i+1}, \ldots, \boldsymbol{v}_n)$$

is a linear transformation. In other words, a multilinear map is one which is linear in each variable. When $n = 2$, we say a multilinear map is **bilinear**. Of course, if $n = 1$ then a multilinear map is simply just linear.

Example 5.2.2. The dot product $\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is bilinear since for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ and $c \in \mathbb{R}$ it holds

$$(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}, (c\boldsymbol{u}) \cdot \boldsymbol{v} = c(\boldsymbol{u} \cdot \boldsymbol{v}), \boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}, \boldsymbol{u} \cdot (c\boldsymbol{v}) = c(\boldsymbol{u} \cdot \boldsymbol{v}).$$

Thus, the dot product is linear in the first and the second factor.

Likewise, the tensor product $\otimes : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^{n\times n}$ is also bilinear since for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ and $c \in \mathbb{R}$ it holds

$$(\boldsymbol{u} + \boldsymbol{v}) \otimes \boldsymbol{w} = \boldsymbol{u} \otimes \boldsymbol{w} + \boldsymbol{v} \otimes \boldsymbol{w}, (c\boldsymbol{u}) \otimes \boldsymbol{v} = c(\boldsymbol{u} \otimes \boldsymbol{v}), \boldsymbol{u} \otimes (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \otimes \boldsymbol{v} + \boldsymbol{u} \otimes \boldsymbol{w}, \boldsymbol{u} \otimes (c\boldsymbol{v}) = c(\boldsymbol{u} \otimes \boldsymbol{v}). \quad \blacksquare$$

Determinants are NOT linear transformations ($\det(A + B) \neq \det(A) + \det(B)$). The determinant map $\det : F^{n\times n} \to F$ is multilinear with respect to both its rows and columns since $\det(A) = \det(A^\top)$. In particular, if A, B, and C are $n \times n$ matrices that differ only in a single row, say the rth row, and assume the rth row of C is obtained by adding corresponding entries in the rth rows of A and B. Then

$$\det(C) = \det(A) + \det(B).$$

Likewise, if A and B are matrices that differ only in a single row, say the rth row, and assume the rth row of B is obtained by scaling the corresponding entries in the rth row of A by $c \in F$. Then $\det(B) = c \det(A)$.

Example 5.2.3. Note that

$$\begin{vmatrix} 4 & 5 & 0 \\ 3 & -1 & 2 \\ 1+0 & 2+4 & 3-2 \end{vmatrix} = \begin{vmatrix} 4 & 5 & 0 \\ 3 & -1 & 2 \\ 1 & 2 & 3 \end{vmatrix} + \begin{vmatrix} 4 & 5 & 0 \\ 3 & -1 & 2 \\ 0 & 4 & -2 \end{vmatrix} \quad \text{and} \quad \begin{vmatrix} 1 & 2 & 3 \\ 4 & 6 & 8 \\ 0 & 1 & 3 \end{vmatrix} = 2 \begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 0 & 1 & 3 \end{vmatrix}. \quad \blacksquare$$

The following is probably the most important property of determinants.

Theorem 5.2.4. *If A and B are $n \times n$ matrices, then* $\det(AB) = \det(A) \det(B)$.

Example 5.2.5. Let $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$. Then $\det A = 6(2) - 1(3) = 9$ and $\det B = 4(2) - 3(1) = 5$. Thus, $\det(A)\det(B) = \boxed{45}$. On the other hand,

$$AB = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix}.$$

Thus, $\det(AB) = 25(13) - 20(14) = 325 - 280 = \boxed{45}$. ∎

Corollary 5.2.6. *A square matrix A is nonsingular if and only if $\det A \neq 0$. In this case, $\det(A^{-1}) = \dfrac{1}{\det(A)}$.*

The previous corollary tells us that the rows or columns are A are linearly dependent, then $\det(A) = 0$. In particular, if A has a repeated row (or column) or a row (or column) of zeros, then $\det(A) = 0$ without further calculation necessary.

For a square matrix A, let E and B be square matrices such that $A = EB$ and E is an elementary matrix. Thus, $\det(A) = \det(E)\det(B)$. If E is a replacement elementary matrix, $\det(E) = 1$ since it is unit triangular. If E is a scaling elementary matrix by a factor of c, then $\det(E) = c$ since it is diagonal. If E is an interchange elementary matrix, then $\det(E) = -1$.

Theorem 5.2.7. *Let A be a square matrix.*

(i) (Replacement) If a multiple of one row of A is added to another row to produce a matrix B, then $\det B = \det A$.

(ii) (Scaling) If one row of A is multiplied by c to produce B, then $\det B = c \cdot \det A$.

(iii) (Interchange) If two rows of A are interchanged to produce B, then $\det B = -\det A$.

Example 5.2.8. Compute $\det A$, where $A = \begin{bmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{bmatrix}$.

Using Theorem 3, we row reduce A to an echelon form in order to compute $\det A$.

$$\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix} = -1(3)(-5) = \boxed{15}.$$ ∎

Example 5.2.9. Compute $\det A$, where $A = \begin{bmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{bmatrix}$.

Again, we row reduce to calculate the determinant.

$$\begin{vmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix} = 2\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

$$= 4\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & -3 & 2 \end{vmatrix} = 4\begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -3 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 4(1)(3)(-3)(1) = \boxed{-36}.$$

■

Example 5.2.10. Compute $\begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{vmatrix}$.

This time we combine methods of cofactors and row reduction.

$$\begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ -2 & -5 & 4 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 1 & 2 & -1 \\ 2 & 5 & -7 & 3 \\ 0 & 3 & 6 & 2 \\ 0 & 0 & -3 & 1 \end{vmatrix} \quad \text{(add Row 2 to Row 4)}$$

$$= -2\begin{vmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ 0 & -3 & 1 \end{vmatrix} \quad \text{(cofactor expand across Column 1)}$$

$$= -2\begin{vmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & -3 & 1 \end{vmatrix} \quad \text{(add -3 Row 1 to Row 2)}$$

$$= 2\begin{vmatrix} 1 & 2 & -1 \\ 0 & -3 & 1 \\ 0 & 0 & 5 \end{vmatrix} \quad \text{(interchange Rows 2 and 3)}$$

$$= \quad 2(-3)(5) = \boxed{-30}.$$

■

Exercises

(Go to Solutions)

For Exercises 1-8, suppose that $\begin{vmatrix} a & b & c \\ e & f & g \\ h & i & j \end{vmatrix} = 2$, $\begin{vmatrix} a & b & c \\ e & f & g \\ x & y & z \end{vmatrix} = -3$, and $\begin{vmatrix} a & b & u \\ e & f & v \\ h & i & w \end{vmatrix} = 5$. Compute the given expression.

1. $\begin{vmatrix} a & u & b \\ e & v & f \\ h & w & i \end{vmatrix}$

♠ 2. $\begin{vmatrix} a & b & c \\ h & i & j \\ e & f & g \end{vmatrix}$

♠ 3. $\begin{vmatrix} 5a & 5b & 5c \\ e & f & g \\ h & i & j \end{vmatrix}$

4. $\begin{vmatrix} 2u & b & a \\ 2v & f & e \\ 2w & i & j \end{vmatrix}$

5. $\begin{vmatrix} a & b & c+u \\ e & f & g+v \\ h & i & j+w \end{vmatrix}$

♠ 6. $\begin{vmatrix} 3a & 3c & 3b \\ h & j & i \\ e+5a & g+5c & f+5b \end{vmatrix}$

♠ 7. $\begin{vmatrix} e & f & g \\ 4a & 4b & 4c \\ h+5x & i+5y & j+5z \end{vmatrix}$

8. $\begin{vmatrix} 2b & a & c+5u \\ 2f & e & g+5v \\ 2i & h & j+5w \end{vmatrix}$

For Exercises 9-19, compute the determinant of the each of the following matrices using row reduction.

9. $\begin{vmatrix} 6 & 1 \\ 2 & 9 \end{vmatrix}$

10. $\begin{vmatrix} 4+i & 3+i \\ -i & -2 \end{vmatrix}$

♠ 11. $\begin{vmatrix} -3 & 6 & -1 \\ -8 & 11 & -3 \\ 2 & -3 & 1 \end{vmatrix}$

12. $\begin{vmatrix} 1 & 2 & 3 \\ -3 & -2 & -4 \\ 5 & 10 & 21 \end{vmatrix}$

13. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 2 & 4 \end{vmatrix}$

14. $\begin{vmatrix} 2 & 26 & 36 \\ 6 & 87 & 113 \\ 5 & 65 & 97 \end{vmatrix}$

♠ 15. $\begin{vmatrix} 2 & 11 & 4 \\ 9 & 1 & 4 \\ 6 & 6 & 5 \end{vmatrix} \pmod{13}$

16. $\begin{vmatrix} 2 & 4 & 7 \\ -2 & -1 & 1 \\ 3 & 6 & 15 \end{vmatrix}$

♠ 17. $\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 3 & 2 & -1 \\ 3 & 5 & 0 & 1 \\ 4 & -1 & -1 & 0 \end{vmatrix} \pmod{7}$

♠ 18. $\begin{vmatrix} 1 & 0 & 1 & 2 & 3 \\ -1 & 1 & 6 & 3 & 1 \\ 0 & 2 & 6 & 3 & 4 \\ -2 & 4 & 3 & 5 & 0 \\ 0 & 3 & 1 & 2 & 1 \end{vmatrix}$

19. $\begin{vmatrix} 2 & 3 & 9 & 1 & 4 \\ 7 & 2 & 2 & 9 & 3 \\ 0 & 0 & 3 & 3 & 0 \\ 5 & 2 & 2 & 7 & 0 \\ 6 & 2 & 6 & 5 & 4 \end{vmatrix}$

For Exercises 20-24, suppose A is a 5×5 real matrix such that $\det(A) = 3$.

♠ 20. Compute $\text{rank}(A)$. Compute $\text{nullity}(A)$.

♠ 21. Compute $\text{corank}(A)$. Compute $\text{conullity}(A)$.

♠ 22. For any $\boldsymbol{b} \in \mathbb{R}^5$, is the linear system $A\boldsymbol{x} = \boldsymbol{b}$ consistent? How many free variables are in the linear system $A\boldsymbol{x} = \boldsymbol{b}$? How many solutions does $A\boldsymbol{x} = \boldsymbol{b}$ have?

♠ 23. Compute the row reduced echelon form of A.

♠ 24. Let $T : \mathbb{R}^5 \to \mathbb{R}^5$ be the linear transformation with standard matrix A. Compute $\ker(T)$. Compute $\text{im}(T)$. Is T one-to-one? Is T onto?

25. Prove Theorem 5.2.4 in the special case that A and B are 2×2 matrices.

"Desires dictate our priorities, priorities shape our choices, and choices determine our actions. The desires we act on determine our changing, our achieving, and our becoming." – Dallin H. Oaks

5.3 Cramer's Rule

Definition 5.3.1. Let $A = \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{a}_n \end{bmatrix}$ be an $n \times n$ matrix and let $\boldsymbol{b} \in F^n$. Then

$$A_i(\boldsymbol{b}) = \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{b} & \dots & \boldsymbol{a}_n \end{bmatrix},$$

where $\boldsymbol{b}$ replaced the ith column vector of A.

Theorem 5.3.2 (Cramer's Rule)**.** *Let A be a $n \times n$ nonsingular matrix. For any $b \in F^n$, the unique solution $\boldsymbol{x}$ of $A\boldsymbol{x} = \boldsymbol{b}$ has entries given by*

$$x_i = \frac{\det A_i(\boldsymbol{b})}{\det A}.$$

Proof. Let I denote the $n \times n$ identity matrix, with column vectors $\boldsymbol{e}_i$. If $A\boldsymbol{x} = \boldsymbol{b}$, then

$$\begin{aligned} A \cdot I_i(\boldsymbol{x}) &= A \begin{bmatrix} \boldsymbol{e}_1 & \dots & \boldsymbol{x} & \dots & \boldsymbol{e}_n \end{bmatrix} = \begin{bmatrix} A\boldsymbol{e}_1 & \dots & A\boldsymbol{x} & \dots & A\boldsymbol{e}_n \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{a}_1 & \dots & \boldsymbol{b} & \dots & \boldsymbol{a}_n \end{bmatrix} = A_i(\boldsymbol{b}). \end{aligned}$$

Therefore,

$$\det(A)\det(I_i(\boldsymbol{x})) = \det(A \cdot I_i(\boldsymbol{x})) = \det A_i(\boldsymbol{b}).$$

But $\det I_i(\boldsymbol{x}) = x_i$, by row reduction along the ith column of $I_i(\boldsymbol{x})$. Therefore,

$$\det(A) \cdot x_i = \det A_i(\boldsymbol{b}),$$

which finishes the proof. □

Note that $\det(A)x_i = \det A_i(\boldsymbol{b})$ holds even if $\det A = 0$.

Example 5.3.3. Use Cramer's rule to solve the system

$$\begin{cases} 3x_1 - 2x_2 = 6 \\ -5x_1 + 4x_2 = 8. \end{cases}$$

Let $A = \begin{bmatrix} 3 & -2 \\ -5 & 4 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 6 \\ 8 \end{bmatrix}$. Thus,

$$\det A = \begin{vmatrix} 3 & -2 \\ -5 & 4 \end{vmatrix} = 12-10 = 2, \quad \det A_1(\boldsymbol{b}) = \begin{vmatrix} 6 & -2 \\ 8 & 4 \end{vmatrix} = 24+16 = 40, \quad \det A_2(\boldsymbol{b}) = \begin{vmatrix} 3 & 6 \\ -5 & 8 \end{vmatrix} = 24+30 = 54.$$

Then by Cramer's rule, the unique solution is given as

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \det A_1(\boldsymbol{b})/\det A \\ \det A_2(\boldsymbol{b})/\det A \end{bmatrix} = \begin{bmatrix} 40/2 \\ 54/2 \end{bmatrix} = \begin{bmatrix} 20 \\ 27 \end{bmatrix}.$$ ■

Definition 5.3.4. Let A be an $n \times n$ matrix. Then the **adjugate** (or **adjoint**) of A, denoted $\operatorname{adj} A$, is given as

$$\operatorname{adj} A = \left[\; C_{ji} \; \right] = \begin{bmatrix} C_{11} & C_{21} & \dots & C_{n1} \\ C_{12} & C_{22} & \dots & C_{n2} \\ \vdots & \vdots & & \vdots \\ C_{1n} & C_{2n} & \dots & C_{nn} \end{bmatrix}$$

Please note that the adjugate matrix is the *transpose* of the matrix of cofactors.

Theorem 5.3.5. *Let A be an $n \times n$ matrix. Then $A \cdot \operatorname{adj}(A) = \operatorname{adj}(A) \cdot A = \det(A) I_n$. In particular, if A is nonsingular then*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A$$

Proof. The jth columns of A^{-1} is a vector $\boldsymbol{x}$ such that

$$A\boldsymbol{x} = \boldsymbol{e}_j.$$

By Cramer's rule, the ith entry in $\boldsymbol{x}$ is given as

$$x_i = \frac{\det A_i(\boldsymbol{e}_j)}{\det A}.$$

But by cofactor expansion across the ith column, we see that

$$\det A_i(\boldsymbol{e}_j) = (-1)^{i+j} \det A_{ji} = C_{ji}.$$

Therefore, the (i, j) entry in A^{-1} is $\dfrac{\det A_i(\boldsymbol{e}_j)}{\det A} = \dfrac{C_{ji}}{\det A}$, which finishes the proof. □

In particular,

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

Note that is formula holds even if $\det A = 0$.

Example 5.3.6. Find the inverse of the matrix $A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix}$.

We begin by finding the nine cofactors:

$$C_{11} = + \begin{vmatrix} -1 & 1 \\ 4 & -2 \end{vmatrix} = -2, \quad C_{12} = - \begin{vmatrix} 1 & 1 \\ 1 & -2 \end{vmatrix} = 3, \quad C_{13} = + \begin{vmatrix} 1 & -1 \\ 1 & 4 \end{vmatrix} = 5$$

$$C_{21} = - \begin{vmatrix} 1 & 3 \\ 4 & -2 \end{vmatrix} = 14, \quad C_{22} = + \begin{vmatrix} 2 & 3 \\ 1 & -2 \end{vmatrix} = -7, \quad C_{23} = - \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = -7$$

$$C_{31} = + \begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4, \quad C_{32} = - \begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} = 1, \quad C_{33} = + \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -3$$

Therefore,

$$\operatorname{adj} A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix}.$$

To finish, we need to compute $\det A$. We could compute it directly like in the previous sections, but instead we use the observation that $\operatorname{adj} A \cdot A = \det A \cdot I_n$.

$$\operatorname{adj} A \cdot A = \begin{bmatrix} -2 & 14 & 4 \\ 3 & -7 & 1 \\ 5 & -7 & -3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 3 \\ 1 & -1 & 1 \\ 1 & 4 & -2 \end{bmatrix} = \begin{bmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{bmatrix}.$$

Thus, $\det A = 14$ and

$$A^{-1} = \begin{bmatrix} -1/7 & 1 & 2/7 \\ 3/14 & -1/2 & 1/14 \\ 5/14 & -1/2 & -3/14 \end{bmatrix}. \qquad \blacksquare$$

In practice, it is not very practical to solve linear systems or compute inverse via Cramer's rule. The method of row-reduction is generally more efficient. On the other hand, the application of Cramer's rule and adjugate matrices is immeasurable in the theory of linear algebra.

For example, if A is an integer matrix, that is, all of its entries are integers, then its determinant and cofactors will all be integers too. After all, determinants are calculated using addition, subtraction, and multiplication. No division required! Thus, $\operatorname{adj} A$ will be an integer matrix too. Thus, if $\det(A) = \pm 1$, we see that A^{-1} will be an integer matrix as well. This fact is very useful for instructor who want to exercise homework questions with "cute" answers so that students feel happy about their linear algebra homework.

Exercises

(Go to Solutions)

For Exercises 1-5, solve the linear system using Cramer's Rule.

1. $\begin{bmatrix} 6 & 1 \\ 2 & 4 \end{bmatrix} \boldsymbol{x} \equiv \begin{bmatrix} 1 \\ 4 \end{bmatrix} \pmod 7$

♠ 2. $\begin{bmatrix} 1 & 3 \\ 2 & -3 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix} \boldsymbol{x} \equiv \begin{bmatrix} 1 \\ 3 \end{bmatrix} \pmod 5$

4. $\begin{bmatrix} 1 & 3 & 5 \\ 0 & -2 & 3 \\ 3 & 2 & 0 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 2 & 2 & 1 \end{bmatrix} \boldsymbol{x} \equiv \begin{bmatrix} 2 \\ 0 \\ 1 \\ 1 \end{bmatrix} \pmod 3$

For Exercises 6-9, compute the adjugate of the matrix A below. Verify that $A(\operatorname{adj} A) = \det(A) I_n$.

♠ 6. $\begin{bmatrix} 1 & 3 \\ 2 & -3 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 2 & 3 \\ 4 & 2 \end{bmatrix} \pmod 5$

♠ 8. $\begin{bmatrix} 1 & 3 & 5 \\ 0 & -2 & 3 \\ 3 & 2 & 0 \end{bmatrix}$

♠ 9. $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 \\ 2 & 1 & 1 & 2 \\ 0 & 2 & 2 & 1 \end{bmatrix} \pmod 3$

QUICK! For Exercises 10-12, using Theorem 5.3.5, the given matrix A and its adjugate matrix $\operatorname{adj}(A)$, calculate the determinant and inverse matrix of A in LESS THAN 60 SECONDS!

10. $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & 1 \\ 2 & 1 & -5 \end{bmatrix}$, $\operatorname{adj}(A) = \begin{bmatrix} 4 & 13 & 5 \\ -3 & -11 & -4 \\ 1 & 3 & 1 \end{bmatrix}$

11. $A = \begin{bmatrix} -6 & 8 & 13 \\ 6 & -8 & -12 \\ 2 & -2 & -4 \end{bmatrix}$, $\operatorname{adj}(A) = \begin{bmatrix} 8 & 6 & 8 \\ 0 & -2 & 6 \\ 4 & 4 & 0 \end{bmatrix}$

12. $A = \begin{bmatrix} 6 & -2 & 4 \\ -3 & -1 & 1 \\ -4 & 5 & 5 \end{bmatrix}$, $\operatorname{adj}(A) = \begin{bmatrix} -10 & 30 & 2 \\ 11 & 46 & -18 \\ -19 & -22 & -12 \end{bmatrix}$

"You can't cross the sea merely by standing and staring at the water." – Rabindranath Tagore

5.4 Cross Products

Definition 5.4.1. Let $\boldsymbol{u}, \boldsymbol{v} \in F^3$. Then we define the **cross product** $\boldsymbol{u} \times \boldsymbol{v}$ to be the vector in F^3 of the form

$$\boldsymbol{u} \times \boldsymbol{v} = (u_2v_3 - u_3v_2,\ u_3v_1 - u_1v_3,\ u_1v_2 - u_2v_1).$$

In the language of determinants, we can define the cross product as

$$\boldsymbol{u} \times \boldsymbol{v} = \left(\begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix}, -\begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix}, \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \right). \tag{5.4.1}$$

Continuing with this observation, the three coordinates in the cross product are the three cofactors of the determinant $\begin{vmatrix} \boldsymbol{e}_1 & u_1 & v_1 \\ \boldsymbol{e}_2 & u_2 & v_2 \\ \boldsymbol{e}_3 & u_3 & v_3 \end{vmatrix} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$ when expanded across the first column. Thus,

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{e}_1 & u_1 & v_1 \\ \boldsymbol{e}_2 & u_2 & v_2 \\ \boldsymbol{e}_3 & u_3 & v_3 \end{vmatrix} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix} \boldsymbol{e}_1 - \begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix} \boldsymbol{e}_2 + \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \boldsymbol{e}_3.$$

There are important distinctions between the cross product and the dot product worth mentioning. First, the cross product of two vectors is itself a vector in F^3. For this reason, it is sometimes called the *vector product*. Conversely, the dot product of two vectors is a scalar in F. For this reason, the dot product is sometimes called the *scalar product*. Similarly, the outer product is sometimes called the *matrix product* since $\boldsymbol{u} \otimes \boldsymbol{v}$ is a matrix. Second, while the dot product is defined for all vectors in F^n, the cross product is only defined here for vectors in F^3.

Example 5.4.2. Let $\boldsymbol{u} = (1, 2, -2)$ and $\boldsymbol{v} = (3, 0, 1)$. Then their cross product is given as

$$\boldsymbol{u} \times \boldsymbol{v} = (2(1) - (-2)(0), (-2)(3) - 1(1), 1(0) - 2(3)) = (2 - 0, -6 - 1, 0 - 6) = \boxed{(2, -7, -6)}. \qquad \blacksquare$$

Definition 5.4.3. If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, then $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ is called the **scalar triple product** of $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$.

Theorem 5.4.4. *If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, then*

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

Proof.

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) \quad = \quad \boldsymbol{u} \cdot \left(\begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix}, -\begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix}, \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} \right)$$

$$= u_1 \begin{vmatrix} v_2 & w_2 \\ v_3 & w_3 \end{vmatrix} - u_2 \begin{vmatrix} v_1 & w_1 \\ v_3 & w_3 \end{vmatrix} + u_3 \begin{vmatrix} v_1 & w_1 \\ v_2 & w_2 \end{vmatrix} = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix},$$

where the last equality is seen by cofactor-expansion across the first column. □

Corollary 5.4.5. *If $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$, then the area of the parallelogram spanned by $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\|\boldsymbol{u} \times \boldsymbol{v}\|$ and the volume of the parallelepiped spanned by $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ is $|\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})|$.*

Example 5.4.6. Compute the scalar triple product $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w})$ for the vectors $\boldsymbol{u} = (3, -2, -5)$, $\boldsymbol{v} = (1, 4, -4)$, and $w = (0, 3, 2)$.

$$\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = \begin{vmatrix} 3 & 1 & 0 \\ -2 & 4 & 3 \\ -5 & -4 & 2 \end{vmatrix} = 3 \begin{vmatrix} 4 & 3 \\ -4 & 2 \end{vmatrix} - \begin{vmatrix} -2 & 3 \\ -5 & 2 \end{vmatrix} = 3(8 + 12) - (-4 + 15) = 60 - 11 = \boxed{49}. \quad \blacksquare$$

Theorem 5.4.7. *Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^3$ and let $k \in \mathbb{R}$. Then*

(i) $\boldsymbol{u} \cdot (\boldsymbol{u} \times \boldsymbol{v}) = 0$;

(ii) $\boldsymbol{v} \cdot (\boldsymbol{u} \times \boldsymbol{v}) = 0$;

(iii) $\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2$;

(iv) $\|\boldsymbol{u} \times \boldsymbol{v}\| = \|\boldsymbol{u}\| \|\boldsymbol{v}\| \sin\theta$;

(v) $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{u} \cdot \boldsymbol{v})\boldsymbol{w}$;

(vi) $(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w} = (\boldsymbol{u} \cdot \boldsymbol{w})\boldsymbol{v} - (\boldsymbol{v} \cdot \boldsymbol{w})\boldsymbol{u}$;

(vii) $\boldsymbol{u} \times \boldsymbol{0} = \boldsymbol{0} \times \boldsymbol{u} = \boldsymbol{0}$;

(viii) $\boldsymbol{u} \times \boldsymbol{v} = -(\boldsymbol{v} \times \boldsymbol{u})$;

(ix) $\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \times \boldsymbol{v} + \boldsymbol{u} \times \boldsymbol{w}$;

(x) $(\boldsymbol{u} + \boldsymbol{v}) \times \boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{w} + \boldsymbol{v} \times \boldsymbol{w}$;

(xi) $k(\boldsymbol{u} \times \boldsymbol{v}) = (k\boldsymbol{u}) \times \boldsymbol{v} = \boldsymbol{u} \times (k\boldsymbol{v})$;

(xii) $\boldsymbol{u} \times \boldsymbol{u} = \boldsymbol{0}$;

There are a few important observations to make. First, the cross product is noncommutative, that is, $\boldsymbol{u} \times \boldsymbol{v} \neq \boldsymbol{v} \times \boldsymbol{u}$. It is instead what we call *anti-commutative*. Second, the cross product is nonassociative, that is, $\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) \neq (\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{w}$. Instead, cross products satisfy what is known as the *Jacobi identity*:

$$\boldsymbol{u} \times (\boldsymbol{v} \times \boldsymbol{w}) + \boldsymbol{v} \times (\boldsymbol{w} \times \boldsymbol{u}) + \boldsymbol{w} \times (\boldsymbol{u} \times \boldsymbol{v}) = \boldsymbol{0}.^{\dagger}$$

Finally, the cross product is necessarily orthogonal to the two factor vectors. This provides a usual tool for creating normal vectors in F^3.

We have seen before that an m-flat in F^n can be constructed in two ways:
Bottom-Up—Using m linearly independent spanning vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$, we construct a *vector equation* for which the flat is the solution set to the vector equation

$$\boldsymbol{x} = \boldsymbol{x}_0 + c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_m\boldsymbol{v}_m,$$

where $\boldsymbol{x}_0$ is a vector on the flat, $c_1, c_2, \ldots, c_m \in F$ are scalar parameters. Of course, a vector equation corresponds to a linear system, which implies the flat is the solution set of this linear system.

Top-Down—Using $n - m$ linearly independent normal vectors $\boldsymbol{n}_1, \ldots, \boldsymbol{n}_{n-m}$, we construct $n - m$ linearly independent *scalar equations* of the form $\boldsymbol{n}_i \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = 0$ or $\boldsymbol{n}_i \cdot \boldsymbol{x} = \boldsymbol{n}_i \cdot \boldsymbol{x}_0$. If $\boldsymbol{n}_i = (a_1, a_2, \ldots, a_n) \in F^n$, then this scalar equation has the form

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = d,$$

for some $d \in F$. Then the flat is the solution to the system of all these scalar equations. These normal vectors, of course, belong to the orthogonal complement of the flat.

How does one transition between these two representations of the flats? If the *Top-Down* representation is given, one could solve the linear system $A\boldsymbol{x} = \boldsymbol{b}$ to find $\boldsymbol{x} = \boldsymbol{x}_0 + c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_m\boldsymbol{v}_m$, where $\boldsymbol{x}_0$ is a particular solution and $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ is a basis for $\text{Nul}(A)$. This gives the *Bottom-Up* representation of the flat.

If we start with the *Bottom-Up* representation $\boldsymbol{x} = \boldsymbol{x}_0 + c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_m\boldsymbol{v}_m$, we need to a matrix A and vector $\boldsymbol{b}$ such that $\text{Nul}(A) = \text{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. As $\text{Nul}(A)^\perp = \text{Row}(A)$, we need to find vector orthogonal to the spanning vectors. This can be best accomplished by creating a matrix B such that $\text{Row}(B) = \text{Nul}(A) = \text{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$. Then $\text{Nul}(B) = \text{Row}(B)^\perp = \text{Nul}(A)^\perp = \text{Row}(A)$. With these normal vectors, we can construct the linear system for the flat.

Example 5.4.8. Consider the 2-flat in F^4 given by the vector equation $\boldsymbol{x} = \boldsymbol{x}_0 + s\boldsymbol{u} + t\boldsymbol{v}$ where $\boldsymbol{x}_0 = (1, 2, 3, 4)$, $\boldsymbol{u} = (1, 0, 0, -1)$, and $\boldsymbol{v} = (2, 1, -3, 0)$. Find a linear system $A\boldsymbol{x} = \boldsymbol{x}$ for which this flat is the solution set.

Let $B = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 2 & 1 & -3 & 0 \end{bmatrix}$. Then

$$B \sim \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -3 & -2 \end{bmatrix}. \quad \text{Hence, } \text{Row}(B)^\perp = \text{Nul}(B) = \text{Span}\left\{ \begin{bmatrix} 0 \\ 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Therefore, the flat is the solution to the following linear system:

$$\begin{cases} (0,3,1,0)\cdot(x_1-1, x_2-2, x_3-3, x_4-4) &= 0 \\ (1,2,0,1)\cdot(x_1-1, x_2-2, x_3-3, x_4-4) &= 0 \end{cases} \Rightarrow \begin{cases} 0x_1 + 3x_2 + 1x_3 + 0x_4 &= 0(1)+3(2)+1(3)+0(4) \\ 1x_1 + 2x_2 + 0x_3 + 1x_4 &= 1(1)+2(2)+0(3)+1(4) \end{cases}$$

$$\Rightarrow \begin{cases} 3x_2 + x_3 &= 9 \\ x_1 + 2x_2 + x_4 &= 9 \end{cases}$$

One can easily check that the vectors in the flat above are solutions to this linear system. ∎

Now, when one wants to find the equation of a hyperplane, then the single normal can be found using determinants in a manner similar to (5.4.1). Of course, when $n = 3$, this is just the cross product.

Example 5.4.9. Find an equation for the plane in $\mathbb{R}^3$ which passes through $(1, -2, 1)$, $(-1, 0, 1)$, and $(3, 2, 0)$.

We can construct the equation using a normal vector. First, consider the plane as a vector equation. Let $\boldsymbol{u} = (3, 2, 0) - (1, -2, 1) = (2, 4, -1)$ and $\boldsymbol{v} = (3, 2, 0) - (-1, 0, 1) = (4, 2, -1)$. If $\boldsymbol{x}_0 = (-1, 0, 1)$, then the plane is given as $\boldsymbol{x} = s\boldsymbol{u} + t\boldsymbol{v} + \boldsymbol{x}_0$. As the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ give the "slope" of the plane, we want to find a vector orthogonal to both $\boldsymbol{u}$ and $\boldsymbol{v}$. Such a vector is $\boldsymbol{u} \times \boldsymbol{v}$. Note

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{e}_1 & 2 & 4 \\ \boldsymbol{e}_2 & 4 & 2 \\ \boldsymbol{e}_3 & -1 & -1 \end{vmatrix} = \begin{vmatrix} 4 & 2 \\ -1 & -1 \end{vmatrix} \boldsymbol{e}_1 - \begin{vmatrix} 2 & 4 \\ -1 & -1 \end{vmatrix} \boldsymbol{e}_2 + \begin{vmatrix} 2 & 4 \\ 4 & 2 \end{vmatrix} \boldsymbol{e}_3 = \begin{bmatrix} -4+2 \\ -(-2+4) \\ 4-16 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \\ -12 \end{bmatrix}.$$

Thus, the equation for the plane can be given as $-2(x-(-1))-2(y-0)-12(z-1)=0$
$\Rightarrow \quad -2x-2-2y-12z+12=0 \quad \Rightarrow \quad \boxed{x+y+6z=5}$. ∎

Exercises

(Go to Solutions)

For Exercises 1-3, compute $\boldsymbol{u} \times \boldsymbol{v}$.

♠ 1. $\boldsymbol{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$

♠ 2. $\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 6 \\ -1 \end{bmatrix}$

♠ 3. $\boldsymbol{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}$

For Exercises 4-5, simplify the expression.

♠ 4. $\left(\begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix} \right) \times \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$

♠ 5. $\left(\begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} - 2 \begin{bmatrix} -1 \\ 2 \\ 5 \end{bmatrix} \right) \times \left(2 \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} \right)$

For Exercises 6-8, use normal vectors to construct a system of linear equations so that the given affine set is the solution set to your linear system.

♠ 6. The plane in $\mathbb{R}^3$ which passes through $(1, -2, 1)$, $(-1, 0, 1)$, and $(3, 2, 0)$.

♠ 7. The hyperplane in $\mathbb{R}^4$ which passes through $(2, -1, 3, -1)$, $(3, 0, -2, 2)$, $(1, -2, 0, 2)$, and $(-1, -2, -2, 4)$.

♠ 8. The plane in $\mathbb{R}^4$ which passes through $(1, 2, 3, 4)$, $(2, 3, 0, 1)$, and $(0, 1, 2, 4)$.

9. Prove Theorem 5.4.7 (ix).

10. Prove Theorem 5.4.7 (xii).

[†]The Jacobi Identity in addition to axioms (g)–(k) makes F^3 equipped with the cross product into a special type of vector space known as a **Lie algebra**.

Chapter 6

Eigenvalues

"Try not to become a man of success, but rather try to become a man of value." – Albert Einstein

6.1 Eigenvalues and Eigenvectors

Example 6.1.1. Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\boldsymbol{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\boldsymbol{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Then

$$A\boldsymbol{u} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -5 \\ -1 \end{bmatrix} \quad \text{and} \quad A\boldsymbol{v} = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\boldsymbol{v}.$$

Notice that while multiplication by A sends $\boldsymbol{u}$ off to who-knows-where, multiplication by A only stretches the vector $\boldsymbol{v}$. This chapter will explore more deeply this type of phenomenon. ∎

Definition 6.1.2. Let A be an $n \times n$ matrix. Then a nonzero vector $\boldsymbol{x}$ is an **eigenvector** of A if there is a scalar λ such that

$$A\boldsymbol{x} = \lambda\boldsymbol{x}.$$

In this case, λ is called an **eigenvalue** of A corresponding to $\boldsymbol{x}$.

It is straight forward to check if a vector is an eigenvector.

Example 6.1.3. Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\boldsymbol{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\boldsymbol{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Then

$$A\boldsymbol{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4 \begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\boldsymbol{u}.$$

On the other hand,

$$A\boldsymbol{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda \begin{bmatrix} 3 \\ -2 \end{bmatrix},$$

since

$$\begin{cases} -9\lambda = 3 \\ 11\lambda = -2 \end{cases}$$

has no solution. Therefore, $\boldsymbol{u}$ is an eigenvector of A with eigenvalue $\lambda = -4$. The vector $\boldsymbol{v}$ is not an eigenvector of A. ∎

It can also be checked whether a scalar is an eigenvalue of a matrix, although this requires row reduction.

Example 6.1.4. Let A be the same matrix as the previous example. Is 7 an eigenvalue of A?

Now, suppose that λ is an eigenvalue of A with eigenvector $\boldsymbol{x}$. Then

$$\begin{aligned} A\boldsymbol{x} &= \lambda\boldsymbol{x} \\ A\boldsymbol{x} - \lambda\boldsymbol{x} &= \boldsymbol{0} \\ A\boldsymbol{x} - \lambda I\boldsymbol{x} &= \boldsymbol{0} \\ (A - \lambda I)\boldsymbol{x} &= \boldsymbol{0} \end{aligned}$$

Therefore, $\lambda = 7$ is an eigenvalue if and only if $A - 7I$ has a nontrivial solution to the homogeneous system $(A - 7I)\boldsymbol{x} = \mathbf{0}$. Now,

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}.$$

But,

$$\left[\begin{array}{cc|c} -6 & 6 & 0 \\ 5 & -5 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 5 & -5 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right].$$

Therefore, $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a nontrivial solution to this homogeneous system and hence an eigenvector of A. Note that

$$A\boldsymbol{x} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 7 \end{bmatrix} = 7 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 7\boldsymbol{x}.$$

Therefore, 7 is an eigenvalue of A with eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. ∎

Notice in the last example, that $\boldsymbol{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ was an eigenvector for A with eigenvalue 7. Now, this is not the only eigenvector for 7. In fact, $\boldsymbol{y} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is also an eigenvector. Note that $A\boldsymbol{y} = \begin{bmatrix} 14 \\ 14 \end{bmatrix} = 7\boldsymbol{y}$. Now, $\boldsymbol{y} = 2\boldsymbol{x}$. In fact, any nonzero multiple of $\boldsymbol{x}$ is an eigenvector of A since

$$A(c\boldsymbol{x}) = c(A\boldsymbol{x}) = c(\lambda x) = \lambda(c\boldsymbol{x}).$$

Furthermore, any nontrivial solution to the homogeneous system $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ is an eigenvector. But this is the null space of $(A - \lambda I)$, a subspace of $\mathbb{R}^n$. This leads to the next definition.

Definition 6.1.5. Let A be an $n \times n$ matrix with eigenvalue λ. Then the null space of $(A - \lambda I)$ is called the **eigenspace** of A corresponding to λ. The dimension of the eigenspace is called the **geometric multiplicity** of the eigenvalue λ and corresponds to the nullity of the matrix $A - \lambda I$.

Since eigenspaces are null spaces of a matrix, they are always subspaces of F^n.

Example 6.1.6. Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is $\lambda = 2$. Find a basis for the eigenspace corresponding to $\lambda = 2$.

We form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}.$$

Thus,

$$\begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

This implies that the null space of $A - 2I$ are the solutions to the equation

$$2x_1 - x_2 + 6x_3 = 0.$$

In other words,

$$\text{Nul}(A - 2I) = \text{Span}\left\{ \begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\} = \text{Span}\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

In fact, the eigenspace is 2-dimensional with basis $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix} \right\}$. ■

Theorem 6.1.7. *The eigenvalues of a triangular matrix are the entries on its main diagonal.*

Proof. Let A be a triangular matrix and let λ be the (i, i)-entry of A. Then $A - \lambda I$ is a singular matrix because $\det(A - \lambda I) = 0$. Therefore, $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution, which implies that λ is an eigenvalue of A. □

This idea of determinants will return later in this chapter.

Example 6.1.8. Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. Since both matrices are triangular, the eigenvalues of A are $\lambda = 3, 0, 2$ and the eigenvalues of B are $\lambda = 4, 1$. ■

Notice from the previous example that 4 appeared twice along the diagonal of B. This is a repeated eigenvalue of multiplicity two. This idea of multiplicity will also return later in this chapter.

Exercises

(Go to Solutions)

For Exercises 1-3, determine if the vector is an eigenvector for $A = \begin{bmatrix} 2 & 6 \\ 3 & -1 \end{bmatrix}$. If so, what is the eigenvalue?

1. $\begin{bmatrix} 6 \\ 3 \end{bmatrix}$ 2. $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ 3. $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

For Exercises 4-6, determine if the vector is an eigenvector for $A = \begin{bmatrix} -6 & -21 & -16 \\ 6 & 17 & 12 \\ -5 & -12 & -8 \end{bmatrix}$. If so, what is the eigenvalue?

♠ 4. $\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ ♠ 5. $\begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}$ ♠ 6. $\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$

For Exercises 7-9, determine if the vector is an eigenvector for $A = \begin{bmatrix} 2 & -5 & -2 \\ 2 & -7 & -3 \\ -4 & 14 & 6 \end{bmatrix}$. If so, what is the eigenvalue?

♠ 7. $\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$ ♠ 8. $\begin{bmatrix} -1 \\ -2 \\ 4 \end{bmatrix}$ ♠ 9. $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

For Exercises 10-12, find an eigenvector for the matrix A and eigenvalue λ. Answers may vary.

♠ 10. $\begin{bmatrix} 3 & 1 \\ 0 & -2 \end{bmatrix}$, $\lambda = 3$ ♠ 11. $\begin{bmatrix} 1 & 3 \\ -1 & 5 \end{bmatrix}$, $\lambda = 2$ ♠ 12. $\begin{bmatrix} -2 & -1 \\ 1 & -4 \end{bmatrix}$, $\lambda = -3$

For Exercises 13-15, determine if the number λ if an eigenvalue of $A = \begin{bmatrix} -2 & 4 & -2 & 1 \\ -15 & 29 & -13 & 23 \\ -30 & 44 & -19 & 34 \\ 0 & -4 & 2 & -3 \end{bmatrix}$.

♠ 13. $\lambda = 3$ ♠ 14. $\lambda = -2$ ♠ 15. $\lambda = 2$

For Exercises 16-17, find a basis for the eigenspace for the matrix and each eigenvalue λ listed. Answers may vary. Also, determine the geometric multiplicities of each listed eigenvalue.

♠ 16. $\begin{bmatrix} -7 & 2 & 32 \\ -8 & 1 & 0 \\ -2 & 1 & 13 \end{bmatrix}, \lambda = 5, -3$

♠ 17. $\begin{bmatrix} -8 & -5 & 5 \\ 20 & 12 & -10 \\ 10 & 5 & -3 \end{bmatrix}, \lambda = -3, 2$

For Exercises 18-18, find a basis for the eigenspace for the matrix and each eigenvalue. Answers may vary. Also, determine the geometric multiplicities of each listed eigenvalue.

18. $\begin{bmatrix} 1 & -1 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

19. Let A be an $n \times n$ matrix. Let λ be an eigenvalue of A with associated eigenvector $\boldsymbol{x}$. Show that if m is a positive integer then λ^m is an eigenvalue of A^m. What is the associated eigenvector?

20. Show that if A is idempotent then $\lambda = 0, 1$.

21. Show that if A is nilpotent then $\lambda = 0$.

"There seems to be some perverse human characteristic that likes to make easy things difficult."
– Warren Buffett

6.2 The Characteristic Polynomial

As observed previously, for an $n \times n$ matrix A, a scalar λ is an eigenvalue if and only if $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nontrivial solution if and only if $A - \lambda I$ is singular if and only if $\det(A - \lambda I) = 0$. In this context, determinants will be a valuable tool to compute eigenvalues.

Definition 6.2.1. Let A be an $n \times n$ matrix. Treating λ as a variable, the value $\det(A - \lambda I)$ is a polynomial of degree n, called the **characteristic polynomial** of A. The roots of the characteristic polynomial are exactly the eigenvalues of A since the roots are the solutions to the equation $\det(A - \lambda I) = 0$. The (**algebraic**) **multiplicity** of each eigenvalue is its multiplicity in the characteristic polynomial.

Example 6.2.2. Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$.

Then the characteristic polynomial of A is

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & 6 & -8 \\ 0 & -\lambda & 6 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (3-\lambda)(-\lambda)(2-\lambda) \\ &= (6 - 5\lambda + \lambda^2)(-\lambda) = -6\lambda + 5\lambda^2 - \lambda^3. \end{aligned}$$

Note that the eigenvalues of A are 3, 0, and 2, where all multiplicities are one.

The characteristic polynomial of B is

$$\begin{aligned} \det(B - \lambda I) &= \begin{vmatrix} 4-\lambda & 0 & 0 \\ -2 & 1-\lambda & 0 \\ 5 & 3 & 4-\lambda \end{vmatrix} = (4-\lambda)(1-\lambda)(4-\lambda) \\ &= (16 - 8\lambda + \lambda^2)(1-\lambda) = 16 - 24\lambda + 9\lambda^2 - \lambda^3. \end{aligned}$$

Note that the eigenvalues of B are 4 and 1, where 4 has multiplicity two and 1 has multiplicity one. ∎

Let A be a matrix with eigenvalue λ. Suppose that m is the geometric multiplicity of λ ($\text{nullity}(A - \lambda I)$) and n is the algebraic multiplicity of λ. Then $1 \leq m \leq n$, that is, the algebraic multiplicity is an upper bound for the geometric multiplicity.

Example 6.2.3. Find the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

We can find the eigenvalues of A by computing the characteristic polynomial of A and factoring it.

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix} = (2-\lambda)(-6-\lambda) - 3(3)$$

$$\begin{aligned} &= (-12 + 4\lambda + \lambda^2) - 9 = -21 + 4\lambda + \lambda^2 \\ &= (\lambda + 7)(\lambda - 3) \end{aligned}$$

Therefore, the eigenvalues of A are $\lambda = -7, 3$. ∎

Example 6.2.4. If $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, then the characteristic polynomial of A is

$$\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$$

and the complex eigenvalues of A are $\pm i$. Since

$$A - iI = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix} \sim \begin{bmatrix} i & 1 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix},$$

we have that $\begin{bmatrix} i \\ 1 \end{bmatrix}$ is an eigenvector. Note that

$$A \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

Likewise,

$$A + iI = \begin{bmatrix} i & -1 \\ 1 & i \end{bmatrix} \sim \begin{bmatrix} i & -1 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & i \\ 0 & 0 \end{bmatrix}.$$

Thus, $\begin{bmatrix} -i \\ 1 \end{bmatrix}$ is an eigenvector of A and

$$A \begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -i \end{bmatrix} = -i \begin{bmatrix} -i \\ 1 \end{bmatrix}.$$

In particular, A is diagonalizable with

$$A = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \begin{bmatrix} -i/2 & 1/2 \\ i/2 & 1/2 \end{bmatrix}.$$

∎

Let A be an $n \times n$ matrix with *real* entries. Then $\overline{A\boldsymbol{x}} = \overline{A}\,\overline{\boldsymbol{x}} = A\overline{\boldsymbol{x}}$. If λ is an eigenvalue of A and $\boldsymbol{x}$ is a corresponding eigenvector, then

$$A\overline{\boldsymbol{x}} = \overline{A}\,\overline{\boldsymbol{x}} = \overline{A\boldsymbol{x}} = \overline{\lambda\boldsymbol{x}} = \overline{\lambda}\,\overline{\boldsymbol{x}}.$$

Then the conjugate of $\boldsymbol{x}$ is also an eigenvector of A whose corresponding eigenvalue is the conjugate of λ. We saw this in the last example. In greater generality, the eigenvalues of a matrix of a field F might lie in some *extension field* E such that $F \subseteq E$.

Theorem 6.2.5. *An $n \times n$ matrix A is nonsingular if and only if 0 is not an eigenvalue of A. In particular, if $A\boldsymbol{x} = \lambda\boldsymbol{x}$, then $A^{-1}x = \dfrac{1}{\lambda}\boldsymbol{x}$.*

Example 6.2.6. Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$. Then its eigenvalues are $\lambda = 3, 0, 2$. In particular, 0 is an eigenvalue of A. Even though, $\mathbf{0}$ cannot be an eigenvector, 0 can be an eigenvalue. Notice that the eigenvalues of $\lambda = 0$ are simply the vectors in $\mathrm{Nul}(A)$, that is, the nullity of A is simply the geometric multiplicity of 0 as an eigenvalue of A. In fact,

$$A\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}. \qquad \blacksquare$$

Definition 6.2.7. Let A and B be $n \times n$ matrices. We say that two matrices are **similar** if there exists a nonsingular $n \times n$ matrix P such that $PAP^{-1} = B$.

Example 6.2.8. The matrices $A = \begin{bmatrix} 1 & 0 & -7 \\ 5 & 1 & 2 \\ -4 & 2 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 15 & -18 & -2 \\ 17 & -17 & -4 \\ 7 & -22 & 4 \end{bmatrix}$ are similar since $P = \begin{bmatrix} 1 & -2 & -1 \\ 1 & -1 & 0 \\ 1 & -4 & -2 \end{bmatrix}$ is nonsingular with $P^{-1} = \begin{bmatrix} 2 & 0 & -1 \\ 2 & -1 & -1 \\ -3 & 2 & 1 \end{bmatrix}$ and

$$PAP^{-1} = \begin{bmatrix} 1 & -2 & -1 \\ 1 & -1 & 0 \\ 1 & -4 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 & -7 \\ 5 & 1 & 2 \\ -4 & 2 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 & -1 \\ 2 & -1 & -1 \\ -3 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & -1 \\ 1 & -1 & 0 \\ 1 & -4 & -2 \end{bmatrix}\begin{bmatrix} 23 & -14 & -8 \\ 6 & 3 & -4 \\ -4 & -2 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 15 & -18 & -2 \\ 17 & -17 & -4 \\ 7 & -22 & 4 \end{bmatrix} = B. \qquad \blacksquare$$

Theorem 6.2.9. *If A and B are two similar matrices, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).*

Likewise, similar matrices have the same determinant, rank, nullity, and trace. Thus, if any two matrices differ on one of these *invariants* then they cannot be similar.

Example 6.2.10. The matrices $A = \begin{bmatrix} 1 & 2 \\ 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ are not similar since $\mathrm{tr}(A) = 1 - 2 = -1 \neq 5 = 3 + 2 = \mathrm{tr}(B)$. $\blacksquare$

Exercises

(Go to Solutions)

For Exercises 1-9, find the eigenvalues and bases for the eigenspaces of the matrix A given. Answers may vary.

♠ 1. $\begin{bmatrix} 10 & 6 \\ -12 & -8 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 6 & 1 \\ -1 & 4 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 35 & 48 \\ -24 & -33 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 4 & -5 \\ 1 & 0 \end{bmatrix}$

♠ 5. $\begin{bmatrix} -1 & -5 \\ 4 & 7 \end{bmatrix}$

♠ 6. $\begin{bmatrix} 5 & -2 \\ 1 & 3 \end{bmatrix}$

♠ 7. $\begin{bmatrix} -6 & -2 & -2 \\ 17 & 6 & 5 \\ 7 & 2 & 3 \end{bmatrix}$

♠ 8. $\begin{bmatrix} 17 & 5 & 5 \\ -41 & -12 & -14 \\ -19 & -6 & -4 \end{bmatrix}$

♠ 9. $\begin{bmatrix} 3 & 1 & 1 \\ -2 & -1 & -1 \\ -7 & -2 & -2 \end{bmatrix}$

For Exercises 10-11, explain why the matrices A and B below are NOT similar. (Sure, do it QUICKLY if you want a challenge, UNDER 60 SECONDS!)

♠ 10. $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 1 & -3 \\ 5 & 2 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 2 & -3 \\ 0 & 3 \end{bmatrix}$

"The line of life is a ragged diagonal between duty and desire." – William R. Alger

6.3 Diagonalization

Theorem 6.3.1. *If $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r$ are eigenvectors of an $n \times n$ matrix A which correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$, respectively, then $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_r\}$ is linearly independent.*

Definition 6.3.2. Let A be an $n \times n$ matrix. We say that A is **diagonalizable** if A is similar to a diagonal matrix D, that is, $A = PDP^{-1}$ for some invertible matrix P.

Example 6.3.3. Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Then A is a diagonalizable matrix. Let $P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$ and $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$. Then $P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$ and

$$A = PDP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 10 & 5 \\ -3 & -3 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}.$$

Next, observe that

$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} \\ &= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 2 \cdot 5^2 & 5^2 \\ -3^2 & -3^2 \end{bmatrix} = \begin{bmatrix} 2 \cdot 5^2 - 3^2 & 5^2 - 3^2 \\ -2 \cdot 5^2 + 2 \cdot 3^2) & -5^2 + 2 \cdot 3^2 \end{bmatrix}. \end{aligned}$$

In fact, it follows by induction that

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ -2 \cdot 5^k + 2 \cdot 3^k & -5^k + 2 \cdot 3^k \end{bmatrix}.$$

Amongst other reasons, diagonalization provides an effective method to compute powers of matrices. ∎

From the previous example, note that

$$A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -5 \end{bmatrix} = 5 \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

and

$$A \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 3 \\ -6 \end{bmatrix} = 3 \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

Thus, the column vectors of P are eigenvectors of A! Furthermore, the diagonal entries of D are the eigenvalues of A!

Theorem 6.3.4 (The Diagonalization Theorem)**.** *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In this case, there is a basis of F^n consisting of eigenvectors of A, called an **eigenvector basis** (or **eigenbasis**). If $A = PDP^{-1}$, then the diagonal entries of D are the eigenvalues of A with multiplicity, the columns of P are eigenvectors, and the eigenvectors in P correspond to the eigenvalues in the same column in D .*

Let $\mathcal{B}$ be an eigenbasis of A and let $\mathcal{E}$ be the standard basis of F^n. Then $P = \underset{\mathcal{E}\leftarrow\mathcal{B}}{P}$ and $P^{-1} = \underset{\mathcal{B}\leftarrow\mathcal{E}}{P}$.

Theorem 6.3.5. *An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.*

Example 6.3.6. If possible, diagonalize the matrix $A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$.

We begin by computing the eigenvalues of A, via the characteristic polynomial of A:

$$\begin{aligned}
\det(A-\lambda I) &= \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} -5-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} - 3\begin{vmatrix} -3 & -3 \\ 3 & 1-\lambda \end{vmatrix} + 3\begin{vmatrix} -3 & -5-\lambda \\ 3 & 3 \end{vmatrix} \\
&= (1-\lambda)[(-5-\lambda)(1-\lambda)+9] - 3[-3(1-\lambda)+9] + 3[-9+3(5+\lambda)] \\
&= -(5+\lambda)(1-\lambda)^2 + 9(1-\lambda) + 9(1-\lambda) - 27 - 27 + 9(5+\lambda) \\
&= -(5+\lambda)(1-\lambda)^2 + 18(1-\lambda) - 54 + 9(5+\lambda) \\
&= -(5+\lambda)(1-\lambda)^2 + 18(1-\lambda) - 9(1-\lambda) = -(5+\lambda)(1-\lambda)^2 + 9(1-\lambda) \\
&= (1-\lambda)[(5+\lambda)(-1+\lambda)+9] = (1-\lambda)[4+4\lambda+\lambda^2] \\
&= -(\lambda-1)(\lambda+2)^2
\end{aligned}$$

Therefore, the eigenvalues of A are 1 and -2 (with multiplicity two). Let $D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.

Next, we compute the eigenvectors of A. We begin with $\lambda = 1$.

$$A - I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, $\text{Nul}(A - I) = \text{Span}\left\{\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right\}$. Next, we use $\lambda = -2$.

$$A + 2I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 3 & 3 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Therefore, $\text{Nul}(A+2I) = \text{Span}\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$. Since A has a basis of eigenvectors, A is diagonalizable. Let $P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$.

To finish, we need to compute P^{-1}.

$$\left[\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 1 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & -1 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{array}\right].$$

Therefore, $P^{-1} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}$ and

$$A = PDP^{-1} = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ -1 & -1 & 0 \end{bmatrix}. \quad \blacksquare$$

Note that a matrix is diagonalizable if and only if each geometric multiplicity is equal to its algebraic multiplicity.

Exercises

(Go to Solutions)

For Exercises 1-5, diagonalize the matrix A. Answers may vary.

♠ 1. $\begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 19 & -8 \\ 40 & -17 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 8 & -20 & 10 \\ 5 & -12 & 5 \\ 5 & -10 & 3 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 6 & 10 & 0 & -4 \\ 3 & 5 & 0 & -2 \\ 12 & 24 & -1 & -8 \\ 15 & 30 & 0 & -11 \end{bmatrix}$

5. $\begin{bmatrix} 8 & 20 & 10 \\ 5 & -12 & 5 \\ 5 & 10 & 3 \end{bmatrix}$

"It's always good to take an orthogonal view of something. It develops ideas." – Ken Thompson

6.4 Orthogonal Diagonalization

Symmetric matrices interestingly bring together the theory of eigenvectors and inner products. Recall a real square matrix U is called *orthogonal* if $U^\top U = I$, that is, $U^T = U^{-1}$ and complex matrix U is called *unitary* if $U^* U = I$, that is, $U^* = U^{-1}$.

Theorem 6.4.1. *If A is a symmetric matrix or a Hermitian matrix, then any two eigenvectors with distinct eigenvalues are orthogonal.*

Example 6.4.2. We note that $A = \begin{bmatrix} 6 & -2 & -1 \\ -2 & 6 & -1 \\ -1 & -1 & 5 \end{bmatrix}$ is symmetric. It can also be checked that

$$\lambda_1 = 8 : \boldsymbol{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}; \qquad \lambda_2 = 6 : \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}; \qquad \lambda_3 = 3 : \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

are all eigenvectors of A. Furthermore, $\boldsymbol{v}_1 \cdot \boldsymbol{v}_2 = \boldsymbol{v}_1 \cdot \boldsymbol{v}_3 = \boldsymbol{v}_2 \cdot \boldsymbol{v}_3 = 0$. Normalizing these vectors:

$$\boldsymbol{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \qquad \boldsymbol{u}_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}, \qquad \boldsymbol{u}_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix},$$

we get a diagonalization of A:

$$A = PDP^{-1} = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 8 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

using an orthogonal matrix P. ∎

Definition 6.4.3. We say a real matrix A is **orthogonally diagonalizable** if there exists an orthogonal matrix P and diagonal matrix D such that $A = PDP^\top = PDP^{-1}$. We say a complex matrix A is **unitarily diagonalizable** if there exists an Hermitian matrix P and diagonal matrix D such that $A = PDP^* = PDP^{-1}$.

The matrix in the previous example is orthogonally diagonalizable. In fact, if A is orthogonally diagonalizable, then $A = PDP^\top$ and $A^\top = (PDP^\top)^\top = (P^\top)^\top D^\top P^\top = PD^\top P^\top = PDP^\top$, that is, $A^\top = A$. Thus, A is symmetric. The converse is also true.

Theorem 6.4.4. *A real matrix A is orthogonally diagonalizable if and only if A is symmetric. A complex matrix A is unitarily diagonalizable if and only if A is Hermitian.*

The method of computing an orthogonal diagonalization is the same as any other diagonalization except we require that our basis of eigenvectors be orthonormal. Normalizing the eigenvectors is simple enough. On the other hand, we cannot simply apply the Gram-Schmidt procedure to an eigenbasis, because the end result may not be an eigenbasis. Instead, we must apply the Gram-Schmidt procedure to a basis for each distinct eigenspace. Since different eigenspaces are mutually orthogonal, the union of these orthogonal bases gives an orthogonal eigenbasis.

Example 6.4.5. Let $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$. It can be shown that $\lambda = 7, -2$ are the eigenvalues of A and

$$\text{Nul}(A - 7I) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix}\right\}, \qquad \text{Nul}(A + 2I) = \text{Span}\left\{\begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}\right\}$$

are the eigenspaces. Applying Gram-Schmidt to the eigenspace of $\lambda = 7$, we get the orthogonal basis

$$\text{Nul}(A - 7I) = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}\right\}.$$

Therefore,

$$\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}, \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}\right\}$$

is an orthogonal eigenbasis. After normalizing, the set

$$\left\{\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}, \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}\right\}$$

is an orthonormal eigenbasis. Therefore,

$$A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{18} & -2/3 \\ 0 & 4/\sqrt{18} & -1/3 \\ 1/\sqrt{2} & 1/\sqrt{18} & 2/3 \end{bmatrix} \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{18} & 4/\sqrt{18} & 1/\sqrt{18} \\ -2/3 & -1/3 & 2/3 \end{bmatrix}$$

is an orthogonal diagonalization of A. ∎

In particular, unitary matrices are the complex analogue of orthogonal matrices. In fact, all of the previous theorems about orthogonal matrices remain true when considering unitary matrices. The same can be said for symmetric and Hermitian matrices.

The set of eigenvalues of a matrix A is called the **spectrum** of A.

Theorem 6.4.6 (The Spectral Theorem for Symmetric (Hermitian) Matrices)**.** *An $n \times n$ symmetric (Hermitian) matrix A has the following properties:*

(i) A has n real eigenvalues, counting multiplicities;

(ii) The geometric and algebraic multiplicities of each eigenvalue of A are equal;

(iii) The eigenspaces of A are mutually orthogonal;

(iv) A is orthogonally/unitarily diagonalizable.

Example 6.4.7. Diagonalize the Hermitian matrix $A = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}$.

We begin with its characteristic polynomial:

$$\begin{aligned} \det(\lambda I - A) &= \begin{bmatrix} \lambda-2 & -1-i \\ -1+i & \lambda-3 \end{bmatrix} = (\lambda-2)(\lambda-3) - (-1-i)(-1+i) \\ &= \lambda^2 - 5\lambda + 6 - 1 - 1 = \lambda^2 - 5\lambda + 4 = (\lambda-4)(\lambda-1). \end{aligned}$$

Therefore, the eigenvalues of A are $\lambda = 1,\ 4$. We then investigate its eigenspaces:

$$\lambda = 1;\ \begin{bmatrix} -1 & -1-i \\ -1+i & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 1+i \\ 0 & 0 \end{bmatrix} \qquad \lambda = 4;\ \begin{bmatrix} 2 & -1-i \\ -1+i & 1 \end{bmatrix} \sim \begin{bmatrix} 2 & -1-i \\ 0 & 0 \end{bmatrix}$$

$$\boldsymbol{v}_1 = \begin{bmatrix} -1-i \\ 1 \end{bmatrix};\quad \boldsymbol{p}_1 = \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \qquad \boldsymbol{v}_2 = \begin{bmatrix} \frac{1}{2}(1+i) \\ 1 \end{bmatrix};\quad \boldsymbol{p}_2 = \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} = \begin{bmatrix} \frac{1+i}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

Note that:

$$\begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix} \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -2-2i+1+i \\ -2+3 \end{bmatrix} = \begin{bmatrix} -1-i \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{2}(1+i) \\ 1 \end{bmatrix} = \begin{bmatrix} 1+i+1+i \\ 1+3 \end{bmatrix} = 4 \begin{bmatrix} \frac{1}{2}(1+i) \\ 1 \end{bmatrix}$$

Let $P = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix}$, which is a unitary matrix. Let $D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$. Then

$$A = PDP^* = \begin{bmatrix} \frac{-1-i}{\sqrt{3}} & \frac{1+i}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} \frac{-1+i}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} 2 & 1+i \\ 1-i & 3 \end{bmatrix}.$$

As can be expected, the eigenvalues of A are real despite A being a non-real matrix. ∎

Let A be a symmetric matrix. Then it has an orthogonal diagonalization given as:

$$\begin{aligned} A &= PDP^\top = \begin{bmatrix} \boldsymbol{u}_1 & \dots & \boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^\top \\ \vdots \\ \boldsymbol{u}_n^\top \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{u}_1 & \dots & \lambda_n\boldsymbol{u}_n \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^\top \\ \vdots \\ \boldsymbol{u}_n^\top \end{bmatrix} \\ &= \lambda_1\boldsymbol{u}_1\boldsymbol{u}_1^\top + \ldots + \lambda_n\boldsymbol{u}_n\boldsymbol{u}_n^\top = \lambda_1(\boldsymbol{u}_1 \otimes \boldsymbol{u}_1) + \ldots + \lambda_n(\boldsymbol{u}_n \otimes \boldsymbol{u}_n). \end{aligned}$$

This last line is called a **spectral decomposition** of A. Each of matrices $B_i = \boldsymbol{u}_i\boldsymbol{u}_i^\top = \boldsymbol{u}_i \otimes \boldsymbol{u}_i$, the outer product of $\boldsymbol{u}_i$ with itself, is an $n \times n$ symmetric matrix with rank 1. The range of B_i is $\text{Span}\{\boldsymbol{u}_i\}$. Furthermore, $B_iB_j = 0$ if $i \neq j$ and $B_i^2 = B_i$, since $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ is a orthonormal set. Thus, the B_i's are idempotent and pairwise "orthogonal." When considering complex vectors, the outer product becomes $\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^*$.

Example 6.4.8. The matrix $A = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix}$ is symmetric and has an orthogonal diagonalization given by

$$A = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 8 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Let $P = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$. Then

$$\begin{aligned} \boldsymbol{u}_1 \otimes \boldsymbol{u}_1 = \boldsymbol{u}_1\boldsymbol{u}_1^\top &= \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix}, \\ \boldsymbol{u}_2 \otimes \boldsymbol{u}_2 = \boldsymbol{u}_2\boldsymbol{u}_2^\top &= \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix} \\ 8\boldsymbol{u}_1\boldsymbol{u}_1^\top + 3\boldsymbol{u}_2\boldsymbol{u}_2^\top &= 8\begin{bmatrix} 4/5 & 2/5 \\ 2/5 & 1/5 \end{bmatrix} + 3\begin{bmatrix} 1/5 & -2/5 \\ -2/5 & 4/5 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ 2 & 4 \end{bmatrix} = A. \quad \blacksquare \end{aligned}$$

Exercises

(Go to Solutions)

For Exercises 1-5, compute an orthogonal (or unitary) diagonalization for the matrix A.

♠ 1. $\begin{bmatrix} 17 & 5 \\ 5 & -7 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 5 & 17 & 8 \\ 17 & 5 & 8 \\ 8 & 8 & 14 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 5 & 3 & -3 & 3 \\ 3 & 5 & 3 & -3 \\ -3 & 3 & 5 & 3 \\ 3 & -3 & 3 & 5 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 4 & 1-i \\ 1+i & 5 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 5 & 0 & 0 \\ 0 & -1 & -1+i \\ 0 & -1-i & 0 \end{bmatrix}$

For Exercises 6-8, compute a spectral decomposition for the matrix A.

♠ 6. A from Exercise 1

♠ 7. A from Exercise 2

♠ 8. A from Exercise 3

9. Show that the inverse of an orthogonal matrix is orthogonal.

10. Show that the product of two orthogonal matrices is orthogonal.

11. Show that the determinant of an orthogonal matrix is ± 1.

12. Show that if U is an orthogonal matrix then $\|U\boldsymbol{x}\| = \|\boldsymbol{x}\|$.

"Transformation literally means going beyond your form." – Wayne Dyer

6.5 Similarity and Linear Transformations

Let V and W be n- and m-dimensional vector spaces with bases $\mathcal{B} = \{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ and $\mathcal{C} = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$, respectively. Let $T : V \to W$ be a linear transformation. Likewise, the coordinate mappings $[\cdot]_{\mathcal{B}} : V \to F^n$ and $[\cdot]_{\mathcal{C}} : W \to F^m$ are linear transformations (in fact, they are isomorphisms). Consider the composite linear transformation $([\cdot]_{\mathcal{C}} \circ T \circ [\cdot]_{\mathcal{B}}^{-1}) : F^n \to F^m$. This is a linear transformation between F^n and F^m. Let A be the standard matrix of this composite transformation. Then

$$[T(\boldsymbol{x})]_{\mathcal{C}} = A[\boldsymbol{x}]_{\mathcal{B}}$$

for all $\boldsymbol{x} \in V$. This matrix $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$ is called the **matrix representation** of T relative to $\mathcal{B}$ and $\mathcal{C}$. In fact,

$$A = \begin{bmatrix} [T(\boldsymbol{b}_1)]_{\mathcal{C}} & [T(\boldsymbol{b}_2)]_{\mathcal{C}} & \ldots & [T(\boldsymbol{b}_n)]_{\mathcal{C}} \end{bmatrix}.$$

The following diagram may be useful in remembering the relationship here:

$$\begin{array}{ccc} \boldsymbol{x} & \xrightarrow{\quad T \quad} & T(\boldsymbol{x}) \\ \downarrow & & \downarrow \\ [\boldsymbol{x}]_{\mathcal{B}} & \xrightarrow{\quad A \quad} & [T(\boldsymbol{x})]_{\mathcal{C}} \end{array}$$

Example 6.5.1. Consider the linear transformation $T : \mathbb{R}^4 \to \mathbb{R}^4$ given by the rule

$$T\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}\right) = \begin{bmatrix} -2x + 11y + z - w \\ 2x - 9y + 2z + w \\ 2x - 4y + 3z \\ 3x + 1.5y + 2z - w \end{bmatrix}.$$

Note that the standard matrix ${}_{\mathcal{E}}[T]_{\mathcal{E}}$ for $T : \mathbb{R}^4 \to \mathbb{R}^4$ is

$${}_{\mathcal{E}}[T]_{\mathcal{E}} = \begin{bmatrix} -2 & 11 & 1 & -1 \\ 2 & -9 & 2 & 1 \\ 2 & -4 & 3 & 0 \\ 3 & 1.5 & 2 & -1 \end{bmatrix}.$$

Suppose $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} \right\}$ is a basis for $V \subseteq \mathbb{R}^4$ and $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 1 \\ 1 \end{bmatrix} \right\}$ is a basis for $W \subseteq \mathbb{R}^4$. Then T restricts to a linear transformation from V to W, that is, $T : V \to W$. Note that

$$T(\boldsymbol{b}_1) = \begin{bmatrix} 19 \\ -6 \\ 3 \\ 8 \end{bmatrix} \qquad T(\boldsymbol{b}_2) = \begin{bmatrix} -3 \\ 0 \\ -1 \\ 1 \end{bmatrix}.$$

In order to compute the matrix representation $T : V \to W$ with respect to $\mathcal{B}$ and $\mathcal{C}$ coordinates, we need to row reduce $\left[\begin{array}{c|c} \mathcal{C} & T(\mathcal{B}) \end{array}\right]$:

$$\left[\begin{array}{c|c} \mathcal{C} & T(\mathcal{B}) \end{array}\right] = \left[\begin{array}{ccc|cc} 1 & 2 & 3 & 19 & -3 \\ -1 & 0 & -2 & -6 & 0 \\ 1 & 0 & 1 & 3 & -1 \\ -1 & 1 & 1 & 8 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 0 & -2 \\ 0 & 1 & 0 & 5 & -2 \\ 0 & 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Thus,

$$[T(\boldsymbol{b}_1)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 5 \\ 3 \end{bmatrix} \qquad \text{and} \qquad [T(\boldsymbol{b}_2)]_{\mathcal{C}} = \begin{bmatrix} -2 \\ -2 \\ 1 \end{bmatrix}.$$

This gives

$$M = \left[\begin{array}{cc} [T(\boldsymbol{b}_1)]_{\mathcal{C}} & [T(\boldsymbol{b}_2)]_{\mathcal{C}} \end{array}\right] = \begin{bmatrix} 0 & -2 \\ 5 & -2 \\ 3 & 1 \end{bmatrix}. \qquad \blacksquare$$

Note that the matrix representation of T depends on the bases chosen for the domain V and codomain W, up to a point.

Theorem 6.5.2. *Let $T : V \to W$ be a linear transformations. Let A and A' be two matrix representations of T. If $m = \dim W$ and $n = \dim V$, then there nonsingular matrices P $(m \times m)$ and Q $(n \times n)$ such that $A = PA'Q^{-1}$.*

Proof. Let $\mathcal{B}, \mathcal{B}'$ be bases for V and $\mathcal{C}, \mathcal{C}'$ be bases for W such that $A = {}_{\mathcal{C}}[T]_{\mathcal{B}}$ and $A' = {}_{\mathcal{C}'}[T]_{\mathcal{B}'}$. Then let $Q = \underset{\mathcal{B} \leftarrow \mathcal{B}'}{P}$ and $P = \underset{\mathcal{C} \leftarrow \mathcal{C}'}{P}$. Let $\boldsymbol{x} \in V$. Then

$$[PA'Q^{-1}[\boldsymbol{x}]_{\mathcal{B}} = PA'[\boldsymbol{x}]_{\mathcal{B}'} = P[T(\boldsymbol{x})]_{\mathcal{C}'} = [T(\boldsymbol{x})]_{\mathcal{C}} = A[\boldsymbol{x}]_{\mathcal{B}}.$$

Thus, $PA'Q^{-1} = A$. □

When $T : V \to V$ is a linear transformation with the same domain and codomain, we often use the same basis for the domain and codomain. We then refer to the **matrix representation** $A = [T]_{\mathcal{B}}$ relative to $\mathcal{B}$. In this case,

$$[T(\boldsymbol{x})]_{\mathcal{B}} = A[\boldsymbol{x}]_{\mathcal{B}}.$$

Corollary 6.5.3. *Let $T : V \to V$ be a linear transformation. Let A and B be two matrix representations of T. Then there exists a nonsingular matrix P such that $A = PBP^{-1}$. In particular, all matrix representations of T are similar to each other.*

Proof. The result follows immediately from Theorem 6.5.2 once it is discovered that $Q = P$. □

Since matrix representations of a linear transformation always remains in the same similarity class, any property invariant on similar matrices can be attached to the linear transformation. For example, if $T : V \to V$ is a linear transformation, then we can define $\det(T)$ to be the determinant of any matrix representation of T. Likewise, we can define $\operatorname{tr}(T)$, $\operatorname{nullity}(T)$, $\operatorname{rank}(T)$, etc. to be the trace, nullity, rank, etc. of any matrix representation of T. This includes the eigenvalues of a matrix.

Example 6.5.4. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be a linear transformation such that

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 5x + 2y \\ 2x + 3y - z \\ 26x + 16y - 2z \end{bmatrix}.$$

Then the standard matrix representation (using the standard basis on $\mathbb{R}^3$) yields

$$A = [T] = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 3 & -1 \\ 26 & 16 & -2 \end{bmatrix}.$$

On the other hand, one could check that $\mathcal{B} = \left\{ \boldsymbol{b}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \boldsymbol{b}_2 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \boldsymbol{b}_3 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \right\}$ is a basis for $\mathbb{R}^3$. Then

$$[T]_{\mathcal{B}} = \begin{bmatrix} [A\boldsymbol{b}_1]_{\mathcal{B}} & [A\boldsymbol{b}_2]_{\mathcal{B}} & [A\boldsymbol{b}_3]_{\mathcal{B}} \end{bmatrix}.$$

To find these coordinate vectors, we solve the linear systems corresponding to the augmented matrix:

$$[\mathcal{B} \mid A\mathcal{B}] = \left[\begin{array}{ccc|ccc} 2 & -1 & -1 & 4 & -1 & -3 \\ -3 & 2 & 1 & -6 & 2 & 3 \\ 1 & 2 & -2 & 2 & 2 & -6 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 3 \end{array}\right]$$

Thus, $[T]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. This is the result of $\mathcal{B}$ being an eigenbasis of A. Thus, the eigenvalues of T are $\lambda = 2, 1, 3$. In $\mathcal{B}$-coordinates, T is just a diagonal matrix, the diagonalization of A. ∎

Exercises

(Go to Solutions)

For Exercises 1-3, given the linear transformation $T : F^n \to F^m$, a basis $\mathcal{B}$ of F^n, and a basis $\mathcal{C}$ of F^m, listed in that order, find the matrix representation ${}_{\mathcal{C}}[T]_{\mathcal{B}}$.

♠ 1. $T : \mathbb{R}^2 \to \mathbb{R}^3$
$T(x, y) = (x + y, 0, 2x + 3y)$,

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\},$$

$$\left\{ \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix} \right\}$$

♠ 2. $T : \mathbb{R}^3 \to \mathbb{R}^2$
$T(x, y, z) = (x + y - 2z, -y + z)$,

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\},$$

$$\left\{ \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix} \right\}$$

♠ 3. $T : \mathbb{Z}_2^4 \to \mathbb{Z}_2$
$T(x, y, z, w) = x + y + z + w \pmod 2$,

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\},$$

$\{1\}$

Appendix A

Notation Index and Frequent Terminology

iff – "if and only if"
$\forall$ – "for all"
$\exists$ – "there exists"
$\{x \mid x$ satisfies some property$\}$ - set builder notation
$\in$ – element of a set
$\subseteq$ – subset
$\leq$ – subspace

Example A.1. $X = \{t(-2,1,3) \mid t \in \mathbb{R}\}$ is the set of all vectors which are real scalar multiples of $(-2,1,3)$. We have that $(0,0,0), (-4,2,6) \in X$, but $(1,0,1) \notin X$. Hence, $\{(0,0,0),(-4,2,6)\} \subseteq X$. As X is a subspace, $X \leq \mathbb{R}^3$. ∎

$\mathbb{N}$ – the set of natural numbers, that is, counting whole numbers, e.g. $0, 1, 2, 3, 4, \ldots$
$\mathbb{Z}$ – the set of integers, that is, whole numbers which are positive, zero, or negative, e.g., $\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$
$\mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\}$ – the set of integers modulo n
$\mathbb{Q}$ – the set of rational numbers, that is, those numbers which can be expressed as a ratio of integers a/b (where $b \neq 0$), e.g. $1/2, 0, -3/5, 7/2, 7/3, 5, \ldots$ Note that a rational number can be written in more than one way.
$\mathbb{R}$ – the set of real numbers, that is, those numbers which are rational or irrational, e.g. π, $\sqrt{2}$, 3.5, 4, -285, 37.4568, $\log_6(3)$, $e^7, \ldots$
$\mathbb{C}$ – the set of complex numbers, that is, those numbers of the form $a + bi$ where $i = \sqrt{-1}$ and a, b are real numbers, e.g., $5, 2 + 4i, -i, 1/2 + 5i/7, \ldots$

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i \qquad \frac{a+bi}{c+di} = \frac{(ac+bd)+(bc-ad)i}{c^2+d^2}$$

Example A.2. $(1+2i)(3+4i) = (3-8)+(4+6)i = \boxed{-5+10i}$,
$\dfrac{1+2i}{3+4i} = \dfrac{(3+8)+(6-4)i}{9+16} = \boxed{\dfrac{11}{25} + \dfrac{2}{25}i}$ ∎

F^n - the vector space of all column vectors with n entries chosen from the field of scalars F
$\mathbf{0}$ - zero vector

$\ker T = \{\boldsymbol{x} \in X \mid T(\boldsymbol{x}) = \mathbf{0}\}$ – kernel of transformation $T : X \to Y$
$\operatorname{im} T = \{T(\boldsymbol{x}) \in Y \mid \boldsymbol{x} \in X\}$ – image (or range) of transformation $T : X \to Y$
one-to-one – $T(\boldsymbol{u}) = T(\boldsymbol{v})$ implies $\boldsymbol{u} = \boldsymbol{v}$ for the transformation $T : X \to Y$. If T is a linear transformation, it is one-to-one iff $\ker T = \{\mathbf{0}\}$.
onto – $\forall \boldsymbol{b} \in Y$, $\exists \boldsymbol{x} \in X$ such that $T(\boldsymbol{x}) = \boldsymbol{b}$ for the transformation $T : X \to Y$, that is, $\operatorname{im} T = Y$.

$m \times n$ matrix – a matrix with m rows (horizontal) by n columns (vertical)
overdetermined system – $m > n$, more rows than columns (tall matrix)

underdetermined system – $m < n$, more columns than rows (fat matrix)
echelon form – pivots form a downward staircase, with zeros below, rows of zeros at bottom
row reduced echelon form (RREF) – pivots are all 1's and form a downward staircase, with zeros below and above, rows of zeros at bottom
replacement – replace a row in a matrix with that row plus a multiple of another row, i.e. $R_i \longrightarrow R_i + cR_j$
Interchange – interchanging or swapping two rows in a matrix, i.e. $R_i \longleftrightarrow R_j$
scaling – multiple a row of a matrix by a scalar, i.e. $R_i \longrightarrow cR_i$

Appendix B

The Substitution and Elimination Methods

Solving linear systems by graphing is very error-prone and problematic, even with the use of technology! Although the geometric approach might be necessary to comprehend solutions to linear systems, algebraic methods prove far superior to solving linear systems. In this appendix, we will discuss two such methods typically presented in a pre-linear algebra setting, beginning with substitution.

The Substitution Method

1. Solve one equation for one of the variables.
2. Substitute the expression for this variable into the other equation.
3. Once you have determined one assignment, back substitute it into the equation from 1.

Example B.1. Solve the system by substitution.

$$\begin{cases} 2x + y = 1 \\ 3x + 4y = 14, \end{cases}$$

We will solve the first equation with respect to y. This gives us $y = 1 - 2x$. We next will substitute this value of y into the second equation, giving:

$$3x + 4y = 14 \quad \Rightarrow \quad 3x + 4(1 - 2x) = 14 \quad \Rightarrow \quad 3x + 4 - 8x = 14 \quad \Rightarrow \quad -5x = 10 \quad \Rightarrow \quad x = -2.$$

We next back substitute $x = -2$ into the above equation. This gives $y = 1 - 2(-2) = 1 + 4 = 5$. Therefore, the solution is $\boxed{(-2, 5)}$. ∎

Example B.2. Solve

$$\begin{cases} x + y - z = -1 \\ 4x - 3y + 2z = 16 \\ 2x - 2y - 3z = 5. \end{cases}$$

We will solve the first equation with respect to z : $z = x + y + 1$. We now substitute this expression into *both* of the remaining equations.

$$\begin{cases} 4x - 3y + 2(x + y + 1) = 16 \\ 2x - 2y - 3(x + y + 1) = 5 \end{cases} \sim \begin{cases} 6x - y + 2 = 16 \\ -x - 5y - 3 = 5 \end{cases} \sim \begin{cases} 6x - y = 14 \\ -x - 5y = 8. \end{cases}$$

We now have to solve a system of 2 equations with 2 unknowns. Again using substitution, notice that $y = 6x - 14$. Plugging this into the last equation gives

$$-x - 5(6x - 14) = 8 \quad \Rightarrow \quad x - 30x + 70 = 8 \quad \Rightarrow \quad -31x = -62 \quad \Rightarrow \quad x = 2.$$

We now substitute this value into the above equation for y and get $y = 6(2) - 14 = 12 - 14 = -2$. Lastly, we substitute both of these values back into above equation for x and get $z = (2) + (-2) + 1 = 1$. Therefore, the solution of the system is $\boxed{(2, -2, 1)}$. ∎

Algebraically, one discovers that a particular system is inconsistent if during the process of solving one runs into a contradiction or a individual solution which has no solution.

Example B.3. Find all solutions for

$$\begin{cases} 8x - 2y = 5 \\ -12x + 3y = 7. \end{cases}$$

We will solve the variable x in the first equation, giving $x = \frac{1}{4}y + \frac{5}{8}$. Substituting this into the second equation gives

$$-12\left(\frac{1}{4}y + \frac{5}{8}\right) + 3y = 7 \quad \Rightarrow \quad -3y - \frac{15}{2} + 3y = 7 \quad \Rightarrow \quad -\frac{15}{2} = 7,$$

which is a contradiction. Therefore, the system is $\boxed{\text{inconsistent}}$ and has no solutions. ∎

The Elimination Method

1. Choose a variable to eliminate and adjust the coefficients of said variable so they are alternating.
2. Add the equations and solve for the remaining variable.
3. Back substitute the variable value into one of the original equations.

Now, we try the elimination method to solve a linear system.

Example B.4. Find all solutions by elimination for

$$\begin{cases} 2x - 3y = 4 \\ 4x + 5y = 3. \end{cases}$$

We will eliminate the variable y. Notice that the coefficients of y are already alternating but we need a common multiple. Between 3 and 5, the least common multiple is $3(5) = 15$. Thus, we will multiply the first equation by 5 and the second by 3. This gives

$$\begin{cases} 10x - 15y = 20 \\ 12x + 15y = 9. \end{cases}$$

If we add the equations, we get $22x = 29 \Rightarrow x = \frac{29}{22}$. Plugging this x value into either equation will give the y-value. Alternatively, we could start the elimination process over again by canceling out the x this time. To do so, we will multiply the first equation by -2 and the second by nothing. This gives

$$\begin{cases} -4x + 6y = -8 \\ 4x + 5y = 3. \end{cases}$$

If we add the equations, we get $11y = -5 \Rightarrow y = -\frac{5}{11}$. Therefore, $\boxed{\left(\frac{29}{22}, -\frac{5}{11}\right)}$ is the solution of the system. ∎

Example B.5. Solve

$$\begin{cases} x - 2y - z = 8 \\ 2x - 3y + z = 23 \\ 4x - 5y + 5z = 53. \end{cases}$$

First, we will eliminate x from the system, which means eliminating x from the first pair of equations and from the second pair of equations. With the first pair, multiply the first equation by -2 : $-2x+4y+2z = -16$. When we add the first 2 equation together, we get $y + 3z = 7$. Next, we eliminate x from the second pair of equations. We will multiply the second equation by -2, which gives $-4x + 6y - 2z = -46$. After we add this equation with the third, we get $3y + 9z = 21$. We now have to solve the 2×2 system

$$\begin{cases} y + 3z = 7 \\ 3y + 9z = 21. \end{cases}$$

But upon further inspection, we notice that the first equation is 3 times the second equation. Therefore, the equations are dependent and the system has infinitely many solutions.

In order to get the general form of the solution, notice that we have $y = 7 - 3z$ and $x = 2y + z + 8 = 2(7 - 3z) + z + 8 = -5z + 22$. Therefore, the general solution has the form $\boxed{(-5z + 22, 7 - 3z, z)}$. ∎

Appendix C

Proofs of Theorems

Theorem 3.2.9 *Two matrices A and B are row equivalent if and only if* $\operatorname{Row} A = \operatorname{Row} B$.

Proof. We will show that if A and B are row equivalent, then $\operatorname{Row} A = \operatorname{Row} B$. Let $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_m\}$ be the rows of A. Certainly, interchange would not change the row space since

$$\operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_m\} = \operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_m\}.$$

Additionally, scaling by a nonzero value c does not affect the span of the rows since

$$c\boldsymbol{r}_i \in \operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_m\} \qquad \text{and} \qquad \boldsymbol{r}_i = \frac{1}{c}(c\boldsymbol{r}_i) \in \operatorname{Span}\{\boldsymbol{r}_1, \ldots, c\boldsymbol{r}_i, \ldots, \boldsymbol{r}_m\}.$$

Finally, replacement also does not change the row space because

$$\boldsymbol{r}_j + c\boldsymbol{r}_i \in \operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_j, \ldots, \boldsymbol{r}_m\} \qquad \text{and} \qquad \boldsymbol{r}_j = (\boldsymbol{r}_j + c\boldsymbol{r}_i) - c\boldsymbol{r}_i \in \operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_i, \ldots, \boldsymbol{r}_j + c\boldsymbol{r}_i, \ldots, \boldsymbol{r}_m\}.$$

Since no row operation changes the row space, $\operatorname{Row} A = \operatorname{Row} B$ for all $A \sim B$.

In the other direction, if the rows of A and B are $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\}$ and $\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$, respectively, then $\operatorname{Span}\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n\} = \operatorname{Span}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}$. Applying Section 2.1 Exercise 20 repeatedly allows us to transform along the following lines (up to relabelling):

$$\operatorname{Span}\{\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n\} = \operatorname{Span}\{\boldsymbol{s}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n\} = \operatorname{Span}\{\boldsymbol{s}_1, \boldsymbol{s}_2, \ldots, \boldsymbol{r}_n\} = \ldots = \operatorname{Span}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n\}.$$

Each of these equalities is obtained by expressing $\boldsymbol{s}_i$ as a linear combination of the current spanning set, which inductively can be decomposed as some sequence of row scalings and row replacements. The relabelling process alluded to above can be accomplished by row interchanges. Therefore, A is row equivalent to B. ∎

Theorem 3.2.10 *If U is an echelon form of the matrix A, then nonzero rows of U form a basis for the row space of A.*

Proof. By Theorem 3.2.9, $\operatorname{Row}(A) = \operatorname{Row}(U)$ since A and U are row equivalent. Thus, a basis for $\operatorname{Row}(U)$ is a basis for $\operatorname{Row}(A)$. If U is in echelon form, then no nonzero row can be written as a linear combination of the other rows (otherwise we could have zeroed out that row). Hence, we have a linearly independent spanning set for $\operatorname{Row}(A)$, which is necessarily a basis. ∎

Theorem 3.3.3 *Let* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. *If $ad - bc \neq 0$, then A is nonsingular with inverse*

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

If $ad - bc = 0$, then A is singular.

Proof. Suppose that $ad - bc \neq 0$. Let $B = \dfrac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Note that

$$AB = \frac{1}{ad-bc}\begin{bmatrix} ad-bc & -ab+ab \\ cd-cd & -bc+ad \end{bmatrix} = I_2.$$

Similarly, $BA = I_2$. Thus, $B = A^{-1}$.

Suppose that $ad - bc = 0$. If $a \neq 0$, then $\dfrac{b}{a}\begin{bmatrix} a \\ c \end{bmatrix} = \begin{bmatrix} b \\ bc/a \end{bmatrix} = \begin{bmatrix} b \\ d \end{bmatrix}$. If $a = 0$, then suppose $b \neq 0$. Similarly, $\dfrac{a}{b}\begin{bmatrix} b \\ d \end{bmatrix} = \begin{bmatrix} a \\ ad/b \end{bmatrix} = \begin{bmatrix} a \\ c \end{bmatrix}$. If $a = b = 0$, then either $\begin{bmatrix} 0 \\ c \end{bmatrix}$ or $\begin{bmatrix} 0 \\ d \end{bmatrix}$ is a multiple of the other. In all cases, one column vector is a multiple of the other. Thus, the columns are linearly dependent, which is singular by the Nonsingular Matrix Theorem. ∎

Theorem 3.3.7 *Let A and B be $n \times n$ invertible matrices, let m be a positive integer, and r is a nonzero real number. Then A^{-1}, AB, $A^\top$, A^m, and rA are also invertible with:*

(i) $(A^{-1})^{-1} = A$ (ii) $(AB)^{-1} = B^{-1}A^{-1}$ (iii) $(A^\top)^{-1} = (A^{-1})^\top$

(iv) $(A^m)^{-1} = (A^{-1})^m := A^{-m}$ (v) $(kA)^{-1} = \dfrac{1}{k}A^{-1}$.

Proof.

(i) Note that if $AA^{-1} = A^{-1}A = I_n$, then A^{-1} has an inverse, A. Since inverses are unique, $(A^{-1})^{-1} = A$.

(ii) Note that
$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = A(I_n)A^{-1} = AA^{-1} = I_n.$$
Similarly, $(B^{-1}A^{-1})(AB) = I_n$. Thus, AB has an inverse and $(AB)^{-1} = B^{-1}A^{-1}$.

(iii) Next, note that
$$(A^{-1})^\top A^\top = (AA^{-1})^\top = I_n^\top = I_n.$$
Similarly, $A^\top(A^{-1})^\top = I_n$. Thus, $A^\top$ has an inverse with $(A^\top)^{-1} = (A^{-1})^\top$.

(iv) Next, note that
$$(A^{-1})^m A^m = \underbrace{A^{-1}\cdots A^{-1}}_{m\text{-times}} \cdot \underbrace{A\cdots A}_{m\text{-times}} = I_n.$$
Similarly, $A^m(A^{-1})^m = I_n$. Thus, $(A^{-1})^m = (A^m)^{-1}$.

(v) Finally, note that
$$\left(\frac{1}{r}A^{-1}\right)(kA) = \left(\frac{1}{k}\cdot k\right)(A^{-1}A) = 1 \cdot I_n = I_n.$$
Similarly, $(kA)\left(\dfrac{1}{k}A^{-1}\right) = I_n$. Thus, $\left(\dfrac{1}{k}A^{-1}\right) = (kA)^{-1}$. ∎

Theorem 4.2.7 *If $S = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_p\} \subseteq F^n$ is an orthogonal set of nonzero vectors, then S is linearly independent.*

Proof. Suppose that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_p\boldsymbol{v}_p = \boldsymbol{0}.$$

Then for each i,

$$\begin{aligned}
\boldsymbol{v}_i \cdot (c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_p\boldsymbol{v}_p) &= \boldsymbol{v}_i \cdot \boldsymbol{0} \\
(\boldsymbol{v}_i \cdot c_i\boldsymbol{v}_1) + (\boldsymbol{v}_i \cdot c_2\boldsymbol{v}_2) + \ldots + (\boldsymbol{v}_i \cdot c_i\boldsymbol{v}_i) + \ldots + (\boldsymbol{v}_i \cdot c_p\boldsymbol{v}_p) &= 0 \\
c_1(\boldsymbol{v}_i \cdot \boldsymbol{v}_1) + c_2(\boldsymbol{v}_i \cdot \boldsymbol{v}_2) + \ldots + c_i(\boldsymbol{v}_i \cdot \boldsymbol{v}_i) + \ldots + c_p(\boldsymbol{v}_i \cdot \boldsymbol{v}_p) &= 0 \\
0 + 0 + \ldots + c_i\|\boldsymbol{v}_i\|^2 + \ldots + 0 &= 0 \\
c_i\|\boldsymbol{v}_i\|^2 &= 0.
\end{aligned}$$

Therefore, $c_i = 0$ or $\|\boldsymbol{v}_i\| = 0$. Since $\boldsymbol{v}_i \neq \boldsymbol{0}$, the latter is impossible. Therefore, $c_i = 0$ for all i, which implies that S is linearly independent. ∎

Theorem 4.2.8 *Let W be a subspace of F^n. Then $W^\perp$ is also a subspace of F^n.*

Proof. Since $\boldsymbol{0}$ is orthogonal to every vector, including those in W, $\boldsymbol{0} \in W^\perp$. Let $\boldsymbol{x}, \boldsymbol{y} \in W^\perp$. Then $\boldsymbol{w} \cdot \boldsymbol{x} = \boldsymbol{w} \cdot \boldsymbol{y} = 0$ for all $\boldsymbol{w} \in W$. Then $\boldsymbol{w} \cdot (\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{w} \cdot \boldsymbol{x} + \boldsymbol{w} \cdot \boldsymbol{y} = 0 + 0 = 0$. Thus, $\boldsymbol{x} + \boldsymbol{y} \in W^\perp$. Finally, let $c \in F$. Then $\boldsymbol{w} \cdot (c\boldsymbol{x}) = c(\boldsymbol{w} \cdot \boldsymbol{x}) = c(0) = 0$ for all $\boldsymbol{w} \in W$. Therefore, $c\boldsymbol{x} \in W^\perp$, which proves the $W^\perp$ is a subspace. ∎

Theorem 4.2.10 *Let A be an $m \times n$ matrix. Then*

$$(\operatorname{Row} A)^\perp = \operatorname{Nul} A.$$

Proof. Let $A = \begin{bmatrix} a_{ij} \end{bmatrix}$. Then $\boldsymbol{a}_i = \begin{bmatrix} \overline{a_{i1}} \\ \overline{a_{i2}} \\ \vdots \\ \overline{a_{in}} \end{bmatrix}$ is a "row" vector of A. Let $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \in \operatorname{Nul} A$. Then

$$\boldsymbol{a}_i \cdot \boldsymbol{x} = a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n = 0.$$

So, $\boldsymbol{x}$ is orthogonal to each row vector. Since inner products are bilinear, this shows that $\boldsymbol{x}$ is orthogonal to the entire row space. ∎

Theorem 4.4.1 *For all $\boldsymbol{u}, \boldsymbol{v} \in F^n$,*

$$|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|.$$

Proof. If $\boldsymbol{u} = \boldsymbol{0}$, then both sides of the inequality are 0 and the inequality holds. Suppose that $\boldsymbol{u} \neq \boldsymbol{0}$. Then

$$\|\operatorname{proj}_{\boldsymbol{u}} \boldsymbol{v}\| = \left\| \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}} \boldsymbol{u} \right\| = \left| \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}} \right| \|\boldsymbol{u}\| = \frac{|\boldsymbol{u} \cdot \boldsymbol{v}|}{\|\boldsymbol{u}\|^2}\|\boldsymbol{u}\| = \frac{|\boldsymbol{u} \cdot \boldsymbol{v}|}{\|\boldsymbol{u}\|}.$$

Also,

$$\|\operatorname{proj}_{\boldsymbol{u}} \boldsymbol{v}\|^2 \leq \|\operatorname{proj}_{\boldsymbol{u}} \boldsymbol{v}\|^2 + \|\boldsymbol{v} - \operatorname{proj}_{\boldsymbol{u}} \boldsymbol{v}\|^2 = \|\boldsymbol{v}\|^2,$$

by the Pythagorean Theorem. Therefore,

$$\|\boldsymbol{v}\| \geq \frac{|\boldsymbol{u} \cdot \boldsymbol{v}|}{\|\boldsymbol{u}\|},$$

which proves the inequality. ∎

Theorem 4.4.2 *For all $\boldsymbol{u}, \boldsymbol{v} \in F^n$,*

$$\|\boldsymbol{u}+\boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|.$$

Proof.

$$\begin{aligned}
\|\boldsymbol{u}+\boldsymbol{v}\|^2 &= (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) = \boldsymbol{u}\cdot\boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{v} + \boldsymbol{v}\cdot\boldsymbol{u} + \boldsymbol{v}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|^2 + 2\mathrm{Re}(\boldsymbol{u}\cdot\boldsymbol{v}) + \|\boldsymbol{v}\|^2 \\
&\leq \|\boldsymbol{u}\|^2 + 2|\boldsymbol{u}\cdot\boldsymbol{v}| + \|\boldsymbol{v}\|^2 \\
&\leq \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2, \qquad \text{by Cauchy-Schwarz Inequality,} \\
&= (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2.
\end{aligned}$$

Taking square roots finishes the proof. ∎

Theorem 4.4.3 *Let $\boldsymbol{u}, \boldsymbol{v} \in F^n$. Then*

$$\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta,$$

where θ is the angle between[†] the two line segments from the origin to the points identified with $\boldsymbol{u}$ and $\boldsymbol{v}$.

Proof. Consider the triangle formed by the vectors $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{u}-\boldsymbol{v}$. Let θ be the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. Then

$$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - 2(\boldsymbol{u}\cdot\boldsymbol{v}) + \|\boldsymbol{v}\|^2.$$

Considering the orthogonal projection $\widehat{\boldsymbol{v}} = \mathrm{proj}_{\boldsymbol{u}}(\boldsymbol{v})$, we have that the triangle formed by $\boldsymbol{v}$, $\widehat{\boldsymbol{v}}$, and $\boldsymbol{v}-\widehat{\boldsymbol{v}}$ is right, with interior angle corresponding to the reference angle of θ, call it $\widehat{\theta}$. Then $\cos\widehat{\theta} = \dfrac{\text{adj}}{\text{hyp}} = \dfrac{\|\widehat{\boldsymbol{v}}\|}{\|\boldsymbol{v}\|}$, or

$$\|\boldsymbol{v}\|\cos\widehat{\theta} = \|\widehat{\boldsymbol{v}}\| = \left|\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{u}\cdot\boldsymbol{u}}\right|\|\boldsymbol{u}\| = \frac{|\boldsymbol{u}\cdot\boldsymbol{v}|}{\|\boldsymbol{u}\|^2}\|\boldsymbol{u}\| = \frac{|\boldsymbol{u}\cdot\boldsymbol{v}|}{\|\boldsymbol{u}\|}.$$

We then conclude that $\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\widehat{\theta} = |\boldsymbol{u}\cdot\boldsymbol{v}|$. Therefore, $\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta$, where the sign will depend if θ is acute, right, or obtuse. ∎

[†]We should mention that the situation is dramatically different between real and complex vector spaces. In the real case, we already have a well-defined notion of angles in the plane. Given that two vectors form a triangle which is always contained in a plane, the notion of angle to agree with the established notion of angle from Trigonometry. We will prove this case below. In the complex case, we do not have a notion of a complex angle lying around. Thus, there are two main schools of thought with regard to complex angles. First, as the field $\mathbb{C}$ can be viewed as a 2-dimensional real vector space, every vector space $\mathbb{C}^n$ can be viewed as a $2n$-dimensional real vector space too. Hence, the angle between the n-complex vectors is simply the real angle between the corresponding $2n$-real vectors. The second idea is to use the Cauchy-Schwarz Inequality. Since $0 \leq |\boldsymbol{u}\cdot\boldsymbol{v}| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$, then $0 \leq \dfrac{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}{|\boldsymbol{u}\cdot\boldsymbol{v}|} \leq 1$. Hence, $\dfrac{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}{|\boldsymbol{u}\cdot\boldsymbol{v}|}$ is a real number inside the range of cosine. Thus, we define θ by the relation $\theta = \cos^{-1}\left(\dfrac{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}{|\boldsymbol{u}\cdot\boldsymbol{v}|}\right)$. This is the meaning we take here of complex angles, and, thus, no proof of the Law of Cosines is necessary here as it is immediate from the definition of "the angle between." We could take this perspective for real angles, which we essentially already have, we choose to prove the Law of Cosines as a theorem instead of using it as a definition so that we do not have the burden of developing Trigonometry from our linear algebraic definition of θ.

Appendix D

Solutions to Exercises

Chapter 1 : Introduction to Linear Algebra

1.1 Linear Systems

1. yes
2. yes
3. no
4. no
5. no
6. no
7. yes
8. no
♠ 9. no
♠ 10. no
♠ 11. yes
♠ 12. no
13. no
14. no
15. no
16. yes
17. no
18. consistent, unique solution
19. consistent, multiple solutions
20. inconsistent
21. consistent, unique solution
22. inconsistent
23. consistent, multiple solutions
24. consistent, unique solution
25. consistent, multiple solutions
26. inconsistent
27. consistent, unique solution
28. homogeneous, $(0, 0)$
29. nonhomogeneous
♠ 30. homogeneous, $(0, 0)$
♠ 31. nonhomogeneous
32. nonhomogeneous
33. nonhomogeneous
34. nonhomogeneous
35. nonhomogeneous
♠ 36. homogeneous, $(0, 0, 0, 0)$
37. nonhomogeneous
38. It is consistent since it is homogeneous. For example, $(0, 0, 0)$ is a solution. This is, in fact, the only solution. A sketch of the three planes shows they intersect at a single point.
♠ 39.

z	(x, y, z)
2	$(7, 5, 2)$
3	$(8, 8, 3)$
-1	$(4, -4, -1)$
-2	$(3, -7, -2)$
π	$(\pi + 5, 3\pi - 1, \pi)$

♠ 40. Hint: Have you done Exercises 28–38 yet?

1.2 Fields

1. Equality of integers is when the numbers are exactly the same. Congruence, on the other hand, is only when the remainders are equal, when divided by p. Additionally, congruence depends on the modulus, that is, two integers can be conguence with respect to one modulus but not another, e.g. $3 \equiv 8 \pmod 5$ but $3 \not\equiv 8 \pmod 7$. In contrast, equality is not relative but absolute.

2. False. *The remainder $a \pmod n$ is read "a modulo n" and $\boldsymbol{n}$ is called the modulus of $\mathbb{Z}_n$.*

3. True

4. False. *Modular division $a/b \pmod n$ is defined by multiplying by the inverse of b with respect to modular **multiplication**.*

5. False. *The set $\mathbb{Z}_n$ is a field with respect to modular addition and multiplication if and only if n is a **prime number**.*

6. False. *If $0 \leq a < n$, then $a \pmod n = \boldsymbol{a}$.*

7. 3 8. 4 9. 0 10. 3 11. 4 12. 2 13. 5

♠ 14. 4 ♠ 15. 2 16. 3 17. 1 ♠ 18. 1 ♠ 19. 11 ♠ 20. 6

21. 2 22. 1 23. 0 24. 3 25. 7 26. $x = -1$

♠ 27. $x = -\frac{5}{6}$ ♠ 28. $x = -\frac{1}{5} + \frac{8}{5}i$ 29. $x = 1$ ♠ 30. $x = 1$ ♠ 31. $x = 3$

♠ 32. $x = 2$ ♠ 33. $x = 5$ 34. $x = 3$ 35. $x = 1$ 36. $x = 2$ 37. $x = 8$ 38. $x = 2$

39. $x = 1$ 40. $x = 1$ 41. $x = 5$ 42. $x = 6$ 43. $x = 3$

♠ 44. Hint: As $\mathbb{Z}_6$ only contains six elements, one could check that the equation has no solution by trial and error. The lack of solution to this equation results from the fact that 2 has no multiplicative inverse modulo 6. In fact, $2(3) \equiv 6 \equiv 0 \pmod 6$. Thus, if 2 did have a reciprocal, then 0 would necessarily have one too. Now, we do not divide by zero. Not ever! The dinosaurs made that mistake and where are they now?

1.3 Vector Spaces

1. False. *For any vector space V, $\mathbf{0} \in V$.*

2. $\begin{bmatrix} 9 \\ 20 \\ 27 \end{bmatrix}$ ♠ 3. $\begin{bmatrix} -9 \\ -8 \\ 5 \end{bmatrix}$ 4. $\begin{bmatrix} 8 \\ 10 \\ 16 \end{bmatrix}$ 5. $\begin{bmatrix} 4 \\ 14 \\ 6 \end{bmatrix}$ 6. $\begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix}$

7. $\begin{bmatrix} 14 \\ 32 \\ 28 \end{bmatrix}$ 8. $\begin{bmatrix} -17 + 26i \\ -14 + 26i \end{bmatrix}$ ♠ 9. $\begin{bmatrix} -8 + 4i \\ 3 + 5i \end{bmatrix}$ 10. $\begin{bmatrix} 1 + 5i \\ 4i \end{bmatrix}$

11. $\begin{bmatrix} -3 + 4i \\ 2 + 19i \end{bmatrix}$ 12. $\begin{bmatrix} -2 + 2i \\ -2 + 8i \end{bmatrix}$ 13. $\begin{bmatrix} 1 - 4i \\ -3 - 14i \\ -5i \end{bmatrix}$ 14. $\begin{bmatrix} 21 + 22i \\ 20 + 12i \\ 24 + 21i \end{bmatrix}$

15. $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ 16. $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ ♠ 17. $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ 18. $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ 19. $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$ 20. $\begin{bmatrix} 1 \\ 4 \\ 4 \end{bmatrix}$

♠ 21. $\begin{bmatrix} 4 \\ 3 \\ 0 \\ 2 \end{bmatrix}$ 22. $\begin{bmatrix} 2 \\ 0 \\ 2 \\ 4 \end{bmatrix}$ 23. $\begin{bmatrix} 2 \\ 2 \\ 2 \\ 1 \end{bmatrix}$

♠ 24. For (d), note $\mathbf{0} = 0\mathbf{0} = (c0)\mathbf{0} = c(0\mathbf{0}) = c\mathbf{0}$.
For (e), $\boldsymbol{u} + (-1)\boldsymbol{u} = 1\boldsymbol{u} + (-1)\boldsymbol{u} = (1-1)\boldsymbol{u} = 0\boldsymbol{u} = \mathbf{0}$. Therefore, $(-1)\boldsymbol{u} = -\boldsymbol{u}$.

♠ 25. Let $a_0 + a_1x + a_2x^2 + \ldots + a_nx^n, b_0 + b_1x + b_2x^2 + \ldots + b_nx^n \in \mathcal{P}_n$. Then $(a_0 + a_1x + a_2x^2 + \ldots + a_nx^n) + (b_0 + b_1x + b_2x^2 + \ldots + b_nx^n) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \ldots + (a_n + b_n)x^n$. Since F is a field, $a_i + b_i \in F$ for all i. Thus, the sum is a vector in $\mathcal{P}_n$. Similarly, if $c \in F$, then $c(a_0 + a_1x + a_2x^2 + \ldots + a_nx^n) = (ca_0) + (ca_1)x + (ca_2)x^2 + \ldots + (ca_n)x^n \in \mathcal{P}_n$ since $ca_i \in F$ for all i. Therefore, $\mathcal{P}_n$ is a vector space.

♠ 26. $0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ over $\mathbb{R}^3$. This violates Theorem 1.3.8 (iii).

♠ 27. $0\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ over $\mathbb{R}^3$. This violates Theorem 1.3.8 (iii).

♠ 28. $1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ over $\mathbb{R}^3$. This violates Axiom *(viii)* from Definition 1.3.1.

♠ 29. $1\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ over $\mathbb{R}^3$. This violates Axiom *(viii)* from Definition 1.3.1.

1.4 Linear Transformations

♠ 1. $T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} (x_1 + x_2) + (y_1 + y_2) \\ 0 \\ 2(x_1 + x_2) + 3(y_1 + y_2) \end{bmatrix} = \begin{bmatrix} (x_1 + y_1) + (x_2 + y_2) \\ 0 \\ (2x_1 + 3y_1) + (2x_2 + 3y_2) \end{bmatrix}$

$= \begin{bmatrix} x_1 + y_1 \\ 2x_1 + 3y_1 \end{bmatrix} + \begin{bmatrix} x_2 + y_2 \\ 2x_2 + 3y_2 \end{bmatrix} = T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + T\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right)$

and

$$T\left(c\begin{bmatrix}x\\y\end{bmatrix}\right)=T\left(\begin{bmatrix}cx\\cy\end{bmatrix}\right)=\begin{bmatrix}cx+cy\\0\\2(cx)+3(cy)\end{bmatrix}=\begin{bmatrix}c(x+y)\\c0\\c(2x+3y)\end{bmatrix}=c\begin{bmatrix}x+y\\0\\2x+3y\end{bmatrix}=cT\left(\begin{bmatrix}x\\y\end{bmatrix}\right).$$

Therefore, T is linear.

♠ 2. $T(1,2)=(3,0,8),\ T(5,-2)=(3,0,4)$ ♠ 3. $\ker T=\{\mathbf{0}\}$

♠ 4. yes, $T(2,-1)=(1,0,1)$

♠ 5. T is one-to-one since kernel is trivial. T is not onto since $(0,1,0)\notin \operatorname{im} T$.

6. $$T\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}+\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)=T\left(\begin{bmatrix}x_1+y_1\\x_2+y_2\\x_3+y_3\end{bmatrix}\right)=\begin{bmatrix}(x_1+y_1)+(x_3+y_3)\\(x_1+y_1)+3(x_2+y_2)\end{bmatrix}=\begin{bmatrix}x_1+y_1+x_3+y_3\\x_1+y_1+3x_2+3y_2\end{bmatrix}$$
$$=\begin{bmatrix}x_1+x_3\\x_1+3x_2\end{bmatrix}+\begin{bmatrix}y_1+y_3\\y_1+3y_2\end{bmatrix}=T\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)+T\left(\begin{bmatrix}y_1\\y_2\\y_3\end{bmatrix}\right)$$
and
$$T\left(c\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right)=T\left(\begin{bmatrix}cx_1\\cx_2\\cx_3\end{bmatrix}\right)=\begin{bmatrix}(cx_1)+(cx_3)\\(cx_1)+3(cx_2)\end{bmatrix}=\begin{bmatrix}c(x_1+x_3)\\c(x_1+3x_2)\end{bmatrix}=c\begin{bmatrix}x_1+x_3\\x_1+3x_2\end{bmatrix}=$$
$$cT\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}\right).$$ Therefore, T is linear.

7. $$T\begin{bmatrix}0\\0\\0\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix},T\begin{bmatrix}1\\1\\2\end{bmatrix}=\begin{bmatrix}3\\4\end{bmatrix}$$

8. $$\ker(T)=\left\{t\begin{bmatrix}3\\-1\\-3\end{bmatrix}\,\middle|\,t\in\mathbb{R}\right\}$$

9. yes, for example $T\begin{bmatrix}1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$. In fact, $T\begin{bmatrix}1+3t\\-t\\-3t\end{bmatrix}=\begin{bmatrix}1\\1\end{bmatrix}$ for any $t\in\mathbb{R}$

10. T is not one-to-one since $\ker(T)\neq\{\mathbf{0}\}$. T is onto since for any $\boldsymbol{b}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$ we have that $T\begin{bmatrix}0\\b_2/3\\b_1\end{bmatrix}=$
$$\begin{bmatrix}0+b_1\\0+3(b_2/3)\end{bmatrix}=\begin{bmatrix}b_1\\b_2\end{bmatrix}$$

11. $T((x_1,y_1,z_1)+(x_2,y_2,z_2)) = T(x_1+x_2, y_1+y_2, z_1+z_2) = \begin{bmatrix} (x_1+x_2)+(y_1+y_2)-2(z_1+z_2) \\ -(y_1+y_2)+(z_1+z_2) \end{bmatrix} = \begin{bmatrix} x_1+x_2+y_1+y_2-2z_1-2z_2 \\ -y_1-y_2+z_1+z_2 \end{bmatrix} = \begin{bmatrix} x_1+y_1-2z_1 \\ -y_1+z_1 \end{bmatrix} + \begin{bmatrix} x_2+y_2-2z_2 \\ -y_2+z_2 \end{bmatrix} = T(x_1,y_1,z_1)+T(x_2,y_2,z_2)$ and $T(c(x,y,z)) = T(cx,xy,cz) = \begin{bmatrix} (cx)+(cy)-2(cz) \\ -(cx)+(cz) \end{bmatrix} = \begin{bmatrix} c(x+y-2z) \\ c(-y+z) \end{bmatrix} = c\begin{bmatrix} x+y-2z \\ -y+z \end{bmatrix} = cT(x,y,z)$. Therefore, T is linear.

♠ 12. $T(1,2,3) = (-3,1)$, $T(1,0,-2) = (5,-2)$

♠ 13. $\ker T = \{t(1,1,1) \mid t \in \mathbb{R}\}$

♠ 14. yes, $T(2,1,0) = (3,-1)$

♠ 15. T is not one-to-one since kernel is nontrivial. T is onto since $\operatorname{im} T = \mathbb{R}^2$. Note that $T(b_1+b_2, -b_2, 0) = (b_1, b_2)$.

16. $T(1,3,2) = (7,6,-6)$

17. yes, $T(1,-1,1) = (1,6,2)$

18. (Answers may vary). Note that $T(2,0,0) = (0,8,0) \neq (0,12,0) = 2(0,6,0) = 2T(1,0,0)$. Thus, scalar multiplication is NOT preserved. Thus, T is not linear.

19. (Answers may vary). Note that $T(1,2,3)+T(0,1,-4) = (2,5)+(2,5) = (4,10) \neq (2,5) = T(1,3,-1) = T[(1,2,3)+(0,1,-4)]$.

20. No, the only vector in the image of T is $(2,5)$.

21. T is not one-to-one since $T(1,2,3) = (2,5) = T(0,1,-4)$ but $(1,2,3) \neq (0,1,-4)$. Likewise, $(0,0) \notin \operatorname{im} T = \{(2,5)\}$.

22. $T((x_1,y_1,z_1,w_1)+(x_2,y_2,z_2,w_2)) = T(x_1+x_2, y_1+y_2, z_1+z_2, w_1+w_2) = (x_1+x_2)+(y_1+y_2)+(z_1+z_2)+(w_1+w_2) = (x_1+y_1+z_1+w_1)+(x_2+y_2+z_2+w_2) = T(x_1,y_1,z_1,w_1)+T(x_2,y_2,z_2,w_2)$ and $T(c(x,y,z,w)) = T(cx,cy,cz,cw) = (cx)+(cy)+(cz)+(cw) = c(x+y+z+w) = cT(x,y,z,w)$. Therefore, T is a linear transformation.

♠ 23. $T(1,0,0,1) \equiv 0$, $T(1,0,1,1) \equiv 1$

♠ 24. $\ker T = \{(0,0,0,0),(1,1,0,0),(1,0,1,0),(1,0,0,1),(0,1,1,0),(0,1,0,1),(0,0,1,1),(1,1,1,1)\}$

♠ 25. yes, $T(1,0,1,1) \equiv 1$

♠ 26. T is not one-to-one since kernel is nontrivial. T is onto since $\operatorname{im} T = \mathbb{Z}_2 = \{0,1\}$.

27. $T(1,2) \equiv 1$, $T(5,2) = 0$

28. $T(1,0,1) \equiv (2,4,1)$, $T(1,2,3) \equiv (3,2,0)$

29. no, the linear system $\begin{cases} x+2y+z \equiv 2 \\ 2x+2y+z \equiv 2 \\ 3x+4y+3z \equiv 3 \end{cases} \pmod 5$ is inconsistent.

30. $T(\boldsymbol{x}+\boldsymbol{y}) = T((x_1,\ldots,x_n)+(y_1,\ldots,y_n)) = T(x_1+y_1,\ldots,x_n+y_n) = a_1(x_1+y_1)+\ldots+a_n(x_n+y_n) = a_1x_1+a_1y_1+\ldots+a_nx_n+a_ny_n = (a_1x_1+\ldots+a_nx_n)+(a_1y_1+\ldots+a_n+y_n) = T(x_1,\ldots,x_n)+T(y_1,\ldots,y_n) = T(\boldsymbol{x})+T(\boldsymbol{y})$ and $T(c\boldsymbol{x}) = T(c(x_1,\ldots,x_n)) = T(cx_1,\ldots,cx_n) = a_1(cx_1)+\ldots+a_n(cx_n) = c(a_1x_1)+\ldots+c(a_nx_n) = c(a_1x_1+\ldots+a_nx_n) = cT(x_1,\ldots,x_n) = cT(\boldsymbol{x})$. Therefore, T is a linear transformation.

31. The transformation T is NOT one-to-one if $n > 1$. To see this, note $T(a_2,-a_1,0,\ldots,0) = a_1a_2 - a_2a_1 + 0 + \ldots + 0 = 0 = T(0,0,\ldots,0)$. If $n = 1$, then T is one-to-one if and only if $a_1 \neq 0$ since $T(x_1) = T(y_1) \quad \Rightarrow \quad a_1x_1 = a_1y_1 \quad \Rightarrow \quad x_1 = y_1$.

32. The transformation T is onto if and only if $a_i \neq 0$ for at least one i. To see this, let $c \in F$. Without the loss of generality, suppose that $a_1 \neq 0$. Then $(c/a_1, 0, \ldots, 0) \in F^n$ and $T(c, 0, \ldots, 0) = a_1(c/a_1) + a_2(0) + \ldots + a_n(0) = c$.

1.5 Augmented Matrices

1. $\left[\begin{array}{cc|c} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \end{array}\right]$
free variables: none

2. $\left[\begin{array}{cccc|c} \boxed{1} & 2 & 1 & 4 & -1 \\ 0 & \boxed{1} & 3 & 5 & 0 \end{array}\right]$
free variables: x_3, x_4

3. $\left[\begin{array}{cccc|c} 0 & \boxed{2} & 1 & 3 & 0 \\ 0 & 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & \boxed{4} & 0 \end{array}\right]$
free variables: x_1

4. $\left[\begin{array}{cccc|c} \boxed{1} & 4 & 2 & 6 & 1 \\ 0 & \boxed{2} & 2 & 1 & 2 \\ 0 & 0 & 0 & \boxed{2} & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$
free variables: x_3

5. $\left[\begin{array}{ccccc|c} \boxed{1} & 3 & 5 & 0 & 1 & 1 \\ 0 & \boxed{3} & 4 & 1 & 2 & 2 \\ 0 & 0 & 0 & \boxed{6} & 1 & 3 \\ 0 & 0 & 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{array}\right]$
free variables: x_3

6. $\left[\begin{array}{ccc} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{array}\right]$
RREF, rank 3

7. $\left[\begin{array}{cccc} \boxed{1} & 2 & 0 & -3 \\ 0 & 0 & \boxed{1} & 5 \end{array}\right]$
RREF, rank 2

8. $\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$
RREF, rank 0

9. $\left[\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$
RREF, rank 0

10. $\left[\begin{array}{cccc} \boxed{3} & -2 & 4 & 5 \end{array}\right]$
echelon form, rank 1

11. $\left[\begin{array}{cccccc} \boxed{1} & 2 & 3 & 4 & 5 & 6 \end{array}\right]$
RREF, rank 1

12. $\left[\begin{array}{ccc} \boxed{1} & 1 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{array}\right]$
echelon form, rank 3

13. $\left[\begin{array}{ccc} 0 & \boxed{1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$
RREF, rank 1

14. $\left[\begin{array}{cccc} \boxed{1} & 2 & 4 & -3 \\ 0 & 0 & \boxed{1} & 5 \end{array}\right]$
echelon form, rank 2

15. $\left[\begin{array}{cccc} \boxed{1} & 0 & 0 & 0 \\ 0 & 0 & \boxed{2} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{array}\right]$
echelon form, rank 3

16. $\left[\begin{array}{cccc} \boxed{1} & 6 & -5 & 0 \\ 0 & 0 & 0 & \boxed{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$
echelon form, rank 2

17. $\left[\begin{array}{cc|c} \boxed{5} & 6 & 3 \\ 0 & \boxed{4} & 12 \end{array}\right]$
echelon form, rank 2

18. $\left[\begin{array}{ccc|c} \boxed{2} & 4 & 3 & 6 \\ 0 & \boxed{2} & 6 & 2 \\ 0 & 0 & \boxed{1} & 4 \end{array}\right]$
echelon form, rank 3

19. NOT echelon form

♠ 20. $\left[\begin{array}{cccc|c} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & \boxed{6} & 7 & 8 & 9 \\ 0 & 0 & 0 & \boxed{3} & 4 \end{array}\right]$
echelon form, rank 3

♠ 21. $\left[\begin{array}{cccc|c} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & \boxed{1} & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$
echelon form, rank 2

♠ 22. NOT echelon form

♠ 23. $\left[\begin{array}{cccc|c} \boxed{1} & 0 & 3 & 0 & 5 \\ 0 & \boxed{1} & 7 & 0 & 9 \\ 0 & 0 & 0 & \boxed{1} & 2 \end{array}\right]$
RREF, rank 3

♠ 24. $\left[\begin{array}{cccc|c} \boxed{1} & 0 & 3 & 4 & 5 \\ 0 & \boxed{1} & 7 & 8 & 9 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$
RREF, rank 2

♠ 25. NOT echelon form

26. NOT echelon form

27. $\left[\begin{array}{ccc|c} \boxed{1} & -\frac{1}{5} & \frac{2}{5} & -\frac{2}{5} \\ 0 & \boxed{2} & 1 & \frac{7}{3} \\ 0 & 0 & 0 & 1 \end{array}\right]$
echelon form, rank 2

28. $\left[\begin{array}{ccc|c} \boxed{1} & 0 & \frac{1}{2} & 0 \\ 0 & \boxed{1} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right]$
RREF, rank 2

29. NOT echelon form

30. $\left[\begin{array}{cccc|c} \boxed{1} & 0 & 0 & 3 & 5 \\ 0 & \boxed{7} & 0 & 0 & 4 \\ 0 & 0 & 0 & \boxed{1} & 6 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$
echelon form, rank 3

31. $\left[\begin{array}{ccc|c} 9 & 4 & 2 & 6 \\ 3 & 5 & 2 & 7 \\ 3 & 17 & 8 & 24 \end{array}\right]$

♠ 32. $\left[\begin{array}{cc|c} 3 & -1 & 0 \\ 0 & -2 & 5 \\ 1 & 7 & 13 \end{array}\right]$

♠ 33. $\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ -2 & -6 & 1 & 12 \end{array}\right]$

34. $\left[\begin{array}{cccc|c} 10 & -7 & 2 & -4 & 10 \\ 3 & 4 & -3 & 1 & 20 \\ 1 & 0 & 4 & 0 & 73 \\ 0 & 6 & -3 & 2 & 10 \end{array}\right]$

35. $\left[\begin{array}{cccc|c} 12 & -2 & 3 & \frac{3}{2} & 13 \\ 1 & 3 & -2 & 0 & 9 \\ 0 & 1 & 17 & -20 & 12 \end{array}\right]$

36. $\begin{cases} x_1 - 2x_2 - 3x_3 = -4 \\ x_2 + 2x_3 = 1 \\ -2x_1 - 2x_3 = 8 \end{cases}$

♠ 37. $\begin{cases} 3x - 2y - z = 4 \\ x + 3z = -3 \\ 0 = 2 \end{cases}$

♠ 38. $\begin{cases} 4x - y - 2z + 5w = 12 \\ -3x + w = 5 \\ x = 0 \end{cases}$

39. $\begin{cases} 12x + z - 9w = 3 \\ 9y + 21z + 25w = -9 \\ 14x - 7y + 36w = 67 \\ 3x + 4y - 7z + 19w = 2 \\ -18x + 2z + 7w = 1 \end{cases}$

40. $\begin{cases} 2x + 2y = 5 \\ 3x + 3y - 3z = 3 \\ -5y + 4z = -1 \end{cases}$
scale Row 2 by 3

♠ 41. $\begin{cases} x - y + 3z = 6 \\ 2y + 5z = 10 \\ 2x + 7y - z = 0 \end{cases}$

♠ 42. $\begin{cases} 3x - y = 0 \\ y = -5/2 \\ x + 7y = 13 \end{cases}$

♠ 43. $\begin{cases} x + y + z = 1 \\ -4y + 3z = 14 \end{cases}$

44. $\begin{cases} 7x + y - 2z \equiv 10 \\ 5x + 5y + 9z \equiv 8 \\ 6x - z \equiv 3 \end{cases} \pmod{11}$

45. $\left[\begin{array}{ccc|c} 1 & 3 & 2 & 3 \\ 0 & 1 & 7 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$

46. $\left[\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & 5 \\ 4 & 4 & 6 & 8 & 17 \end{array}\right]$

47. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & -1 & 2 \\ 0 & 1 & 3 & 5 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$

48. $\left[\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 1 & 2 & 2 & 3 \\ 9 & 4 & 7 & 2 \end{array}\right]$

49. $\left[\begin{array}{ccc|c} 1 & 2 & 3 & 6 \\ 0 & -1 & -2 & -11 \\ 0 & 0 & 0 & 16 \end{array}\right]$

♠ 50. $\left[\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 3 & 2 & 1 & 4 \\ 0 & 0 & 0 & 2 \end{array}\right]$

♠ 51. $\left[\begin{array}{cccc|c} 1 & 1 & 2 & 0 & 3 \\ 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right]$

♠ 52. $\left[\begin{array}{cccc|c} 1 & 0 & 2 & 2 & 1 \\ 0 & 1 & 3 & 1 & 2 \\ 0 & 1 & 2 & 3 & 4 \end{array}\right]$

53. Interchange Rows 1 and 2.

54. Interchange Rows 1 and 3,
replace Row 1 with Row1 − Row2,
scale Row 1 by $\frac{1}{2}$.

55. replace Row 2 with Row2 − 3Row1,
scale Row 2 by $\frac{1}{2}$

56. replace Row 2 with Row2 − 3Row1,
replace Row 3 with Row3 − 2Row1,
replace Row 4 with Row4 + 2Row1

57. replace Row 1 with Row1 − 6Row3,
replace Row 2 with Row2 − 2Row3,
replace Row 4 with Row4 − 3Row3

1.6 Reduction of Linear Systems

1. $\left[\begin{array}{cc} \boxed{1} & 2 \\ 0 & \boxed{13} \end{array}\right] \sim \left[\begin{array}{cc} \boxed{1} & 0 \\ 0 & \boxed{1} \end{array}\right]$,

replace Row 2 with Row 2 + 4Row 1
scale Row 2 by $\frac{1}{13}$,
replace Row 1 with Row 1 − 2Row 2

♠ 2. $\left[\begin{array}{ccc} \boxed{1} & 2 & 3 \\ 0 & \boxed{1} & 3 \\ 0 & 0 & \boxed{1} \end{array}\right] \sim \left[\begin{array}{ccc} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{array}\right]$,

replace Row 2 with Row 2 − 2Row 1,
replace Row 3 with Row 3 − 2Row 1,
scale Row 2 by −1, scale Row 3 by −1
replace Row 2 with Row 2 − 3Row 3,
replace Row 1 with Row 1 − 3Row 3,
replace Row 1 with Row 1 − 2Row 2

♠ 3. $\left[\begin{array}{cccc} \boxed{1} & 2 & 2 & 1 \\ 0 & \boxed{-3} & -6 & -4 \\ 0 & 0 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{cccc} \boxed{1} & 0 & -2 & -\frac{5}{3} \\ 0 & \boxed{1} & 2 & \frac{4}{3} \\ 0 & 0 & 0 & 0 \end{array}\right]$,

replace Row 2 with Row 2 − 2Row 1,
replace Row 3 with Row 3 − Row 1,
replace Row 3 with Row 3 − Row 2,
scale Row 2 by $-\frac{1}{3}$,
replace Row 1 with Row 1 − 2Row 2

♠ 4. $\left[\begin{array}{cccc} \boxed{1} & 5 & 6 & -5 \\ 0 & \boxed{1} & 3 & 3 \\ 0 & 0 & \boxed{25} & 61 \\ 0 & 0 & 0 & \boxed{\frac{712}{25}} \end{array}\right] \sim \left[\begin{array}{cccc} \boxed{1} & 0 & 0 & 0 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & \boxed{1} & 0 \\ 0 & 0 & 0 & \boxed{1} \end{array}\right]$,

replace Row 2 with Row 2 − 3Row 1,
replace Row 4 with Row 4 + 3Row 1,
interchange Rows 2 and 3,
replace Row 3 with Row 3 + 15Row 2,
replace Row 4 with Row 4 − 20Row 2,
replace Row 4 with Row 4 + $\frac{42}{25}$Row 3
scale Row 4 by $\frac{25}{712}$,
replace Row 1 with Row 1 + 5Row 4,
replace Row 2 with Row 2 − 3Row 4,
replace Row 3 with Row 3 − 61Row 4,
scale Row 3 by $\frac{1}{25}$,
replace Row 1 with Row 1 − 6Row 3,
replace Row 2 with Row 2 − 3Row 3,
replace Row 1 with Row 1 − 5Row 2

5. $\left[\begin{array}{ccc|c} \boxed{1} & 2 & 0 & 9 \\ 0 & \boxed{2} & 4 & 6 \\ 0 & 0 & \boxed{5} & 10 \end{array}\right] \sim \left[\begin{array}{ccc|c} \boxed{1} & 0 & 0 & 11 \\ 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & \boxed{1} & 2 \end{array}\right]$,

scale Row 2 by $\frac{1}{2}$,
scale Row 3 by $\frac{1}{5}$,
replace Row 2 with Row 2 $-$ 2Row 3
replace Row 1 with Row 1 $-$ 2Row 2

6. $\left[\begin{array}{ccc|c} \boxed{8} & 16 & 0 & 24 \\ 0 & \boxed{9} & 18 & 9 \\ 0 & 0 & \boxed{4} & 8 \end{array}\right] \sim \left[\begin{array}{ccc|c} \boxed{1} & 0 & 0 & 9 \\ 0 & \boxed{1} & 0 & -3 \\ 0 & 0 & \boxed{1} & 2 \end{array}\right]$,

interchange Rows 2 and 3,
scale Row 1 by $\frac{1}{8}$,
scale Row 2 by $\frac{1}{9}$,
scale Row 3 by $\frac{1}{4}$,
replace Row 2 with Row 2 $-$ 2Row 3
replace Row 1 with Row 1 $-$ 2Row 2

7. $\left[\begin{array}{cccc|c} \boxed{3} & 0 & 6 & 9 & 18 \\ 0 & \boxed{2} & 4 & 0 & 8 \\ 0 & 0 & 0 & \boxed{3} & 6 \end{array}\right] \sim \left[\begin{array}{cccc|c} \boxed{1} & 0 & 2 & 0 & 9 \\ 0 & \boxed{1} & 2 & 0 & -3 \\ 0 & 0 & 0 & \boxed{1} & 2 \end{array}\right]$,

scale Row 1 by $\frac{1}{3}$,
scale Row 2 by $\frac{1}{2}$,
scale Row 3 by $\frac{1}{3}$,
replace Row 1 with Row 1 $-$ 3Row 3

♠ 8. $\left[\begin{array}{cc} \boxed{1} & 2+5i \\ 0 & \boxed{-6-6i} \end{array}\right] \sim \left[\begin{array}{cc} \boxed{1} & 0 \\ 0 & \boxed{1} \end{array}\right]$,

scale Row 1 by $\frac{1}{1+i}$,
replace Row 2 with Row 2 $-$ $(2-i)$Row 1
scale Row 2 by $\frac{1}{-6-6i}$,
replace Row 1 with Row 1 $-$ $(2+5i)$Row 2

9. $\left[\begin{array}{cc} \boxed{1} & 2i \\ 0 & \boxed{4+9i} \end{array}\right] \sim \left[\begin{array}{cc} \boxed{1} & 0 \\ 0 & \boxed{1} \end{array}\right]$,

scale Row 1 by $\frac{1}{6-2i}$,
replace Row 2 with Row 2 $-$ $(-5+i)$Row 1
scale Row 2 by $\frac{1}{4+9i}$,
replace Row 1 with Row 1 $-$ $(2i)$Row 2

♠ 10. $\left[\begin{array}{cc} \boxed{1} & 12 \\ 0 & \boxed{2} \end{array}\right] \sim \left[\begin{array}{cc} \boxed{1} & 0 \\ 0 & \boxed{1} \end{array}\right] \pmod{17}$,

scale Row 1 by $\frac{1}{2} \equiv 9 \pmod{17}$,
replace Row 2 with Row 2 $-$ 5Row 1,
scale Row 2 by $\frac{1}{2} \equiv 9 \pmod{17}$,
replace Row 1 with Row 1 $-$ 12Row 2

♠ 11. $\left[\begin{array}{ccc} \boxed{2} & 1 & 2 \\ 0 & \boxed{5} & 6 \\ 0 & 0 & \boxed{2} \end{array}\right] \sim \left[\begin{array}{ccc} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{array}\right] \pmod{7}$,

interchange Rows 1 and 2,
replace Row 3 with Row 3 $-$ 3Row 1,
replace Row 3 with Row 3 $-$ $\frac{2}{5}$Row 2 $\equiv$ Row 3 $-$ 6Row 2 $\equiv$ Row 3 + Row 2
scale Row 3 by $\frac{1}{2} \equiv 4 \pmod{7}$,
replace Row 2 with Row 2 $-$ 6Row 3,
replace Row 1 with Row 1 $-$ 2Row 3,
scale Row 2 by $\frac{1}{5} \equiv 3 \pmod{7}$,
replace Row 1 with Row 1 $-$ Row 2,
scale Row 1 by $\frac{1}{2} \equiv 4 \pmod{7}$

12. $\begin{bmatrix} \boxed{1} & 3 & 6 \\ 0 & \boxed{1} & 5 \\ 0 & 0 & \boxed{5} \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix} \pmod 7$,

replace Row 2 with Row 2 − 2Row 1,
scale Row 2 by $-1 \equiv 6 \pmod 7$,
replace Row 3 with Row 3 − 5Row 2,
scale Row 3 by $\frac{1}{5} \equiv 3 \pmod 7$,
replace Row 1 with Row 1 − 6Row 3,
replace Row 2 with Row 2 − 5Row 3,
replace Row 1 with Row 1 − 3Row 2

13. $\begin{bmatrix} \boxed{1} & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 6 \\ 0 & 0 & \boxed{5} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} \boxed{1} & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 6 \\ 0 & 0 & \boxed{1} & 5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \pmod 7$,

replace Row 2 with Row 2 − 1Row 1,
replace Row 3 with Row 3 − 3Row 1,
replace Row 4 with Row 4 − 1Row 4,
scale Row 2 by $\frac{1}{2} \equiv 4 \pmod 7$,
replace Row 3 with Row 3 + Row 2,
replace Row 4 with Row 4 − 4Row 2,
interchange Rows 3 and 4,
scale Row 3 by $\frac{1}{5} \equiv 3 \pmod 7$

14. $\left[\begin{array}{ccc|c} 1 & -2 & 3 & 1 \\ 2 & -2 & 5 & 6 \\ 2 & 3 & -4 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -1 & 3 & 1 \\ 0 & 5 & -10 & 0 \\ 0 & 0 & -1 & 4 \end{array}\right]$,
$\boldsymbol{x} = (5, -8, -4)$

15. $\left[\begin{array}{ccc|c} 0 & 2 & 6 & 4 \\ 1 & 3 & 2 & 5 \\ 3 & 3 & 3 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & 2 & 5 \\ 0 & 2 & 6 & 4 \\ 0 & 0 & 15 & 0 \end{array}\right]$,
$\boldsymbol{x} = (-1, 2, 0)$

16. $\left[\begin{array}{ccc|c} 3 & 2 & 1 & 1 \\ 3 & 8 & 3 & 5 \\ 0 & 3 & 1 & 6 \end{array}\right] \sim \left[\begin{array}{ccc|c} 3 & 2 & 1 & 1 \\ 0 & 3 & 1 & 2 \\ 0 & 0 & 0 & 4 \end{array}\right]$,
inconsistent

17. $\left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 0 & 3 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 3 & 2 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right] \pmod 5$
$\boldsymbol{x} \equiv (2, 2, 1)$

18. $\left[\begin{array}{ccccc|c} 1 & 3 & 5 & 4 & 6 & 0 \\ 2 & 4 & 4 & 3 & 1 & 5 \\ 0 & 0 & 0 & 1 & 0 & 5 \\ 6 & 2 & 6 & 3 & 0 & 2 \end{array}\right] \sim \left[\begin{array}{ccccc|c} 1 & 3 & 5 & 4 & 6 & 0 \\ 0 & 5 & 1 & 2 & 3 & 5 \\ 0 & 0 & 3 & 5 & 3 & 4 \\ 0 & 0 & 0 & 1 & 0 & 5 \end{array}\right] \pmod 7$
$\boldsymbol{x} \equiv (4 + 3t, 6 + t, 6t, 5, t)$

19. $\left[\begin{array}{cc|c} 2 & 4 & 0 \\ 1 & 3 & 2 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right] \pmod 5$,
$\boldsymbol{x} \equiv (1, 2)$

20. $\left[\begin{array}{cc|c} 1 & 2 & 2 \\ 0 & -4 & 8 \\ 2 & 2 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 6 \\ 0 & 1 & -2 \\ 0 & 0 & -2 \end{array}\right]$,
inconsistent

♠ 21. $\left[\begin{array}{cc|c} 1 & 1 & 8 \\ 1 & -1 & 4 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 6 \\ 0 & 1 & 2 \end{array}\right]$,
$\boldsymbol{x} = (6, 2)$

♠ 22. $\left[\begin{array}{ccc|c} 1 & -1 & 0 & 6 \\ 2 & 0 & -3 & 16 \\ 0 & 2 & 1 & 4 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{array}\right]$,
$\boldsymbol{x} = (8, 2, 0)$

23. $\left[\begin{array}{ccc|c} 0 & 3 & 6 & 3 \\ 1 & 2 & -3 & 2 \\ 4 & 0 & -3 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}\right]$,
$\boldsymbol{x} = (0, 1, 0)$

24. $\left[\begin{array}{ccc|c} 0 & 3 & -2 & 0 \\ 1 & 1 & 3 & 12 \\ 6 & 2 & 1 & 13 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$,
$\boldsymbol{x} = (1, 2, 3)$

25. $\left[\begin{array}{ccc|c} 1 & 4 & 7 & 3 \\ 2 & 5 & 8 & 3 \\ 3 & 6 & 9 & 3 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$,
$\boldsymbol{x} = (t - 1, -2t + 1, t)$

♠ 26. $\left[\begin{array}{ccc|c} 2 & -2 & -2 & 2 \\ 2 & 3 & 1 & 2 \\ 3 & 2 & 0 & 0 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -2/5 & 1 \\ 0 & 1 & 3/5 & 0 \\ 0 & 0 & 0 & -15 \end{array}\right]$,
inconsistent

27. $\left[\begin{array}{ccc|c} 7 & 8 & 9 & 10 \\ 4 & 5 & 6 & 10 \\ 1 & 2 & 3 & 10 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & -1 & -10 \\ 0 & 1 & 2 & 10 \\ 0 & 0 & 0 & 0 \end{array}\right]$,
$\boldsymbol{x} = (10 + t, 10 - 2t, t)$

28. $\left[\begin{array}{ccc|c} 2 & 1 & -1 & -10 \\ 1 & -3 & -2 & -4 \\ 3 & 1 & 1 & 15 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 11 \end{array}\right]$,
$\boldsymbol{x} = (3, -5, 11)$

29. $\left[\begin{array}{ccc|c} 2 & 1 & 4 & 199 \\ 2 & 10 & 15 & 984 \\ -1 & 0 & 10 & 58 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 42 \\ 0 & 1 & 0 & 75 \\ 0 & 0 & 1 & 10 \end{array}\right]$,
$\boldsymbol{x} = (42, 75, 10)$

30. $\left[\begin{array}{ccc|c} 5 & 2 & 6 & 11 \\ 0 & 1 & 2 & 5 \\ 3 & 2 & 1 & 7 \\ 2 & 1 & 4 & 6 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1/8 \\ 0 & 1 & 0 & 13/4 \\ 0 & 0 & 1 & 7/8 \\ 0 & 0 & 0 & -3/2 \end{array}\right]$,
inconsistent

31. $\left[\begin{array}{ccc|c} 2 & 4 & -2 & 10 \\ 16 & -23 & 5 & 33 \\ 4 & 2 & -6 & 2 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right]$,
$\boldsymbol{x} = (4, 2, 3)$

32. $\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 3 & 1 & 0 & 1 \\ 1 & 1 & 2 & 0 & 0 \\ 2 & 0 & 1 & 4 & 0 \end{array}\right] \sim \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right]$
$\boldsymbol{x} = (3, 1, -2, -1)$

Chapter 2 : The Algebra and Geometry of Vectors

2.1 Vector Equations

1. $x_1 \begin{bmatrix} 1 \\ 7 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 6 \\ 1 \end{bmatrix} + x_3 \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 3 \end{bmatrix}$

♠ 2. $x_1 \begin{bmatrix} 3 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 7 \\ -2 \end{bmatrix} + x_3 \begin{bmatrix} -5 \\ 6 \end{bmatrix} = \begin{bmatrix} -5 \\ 4 \end{bmatrix}$

♠ 3. $x_1 \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 6 \\ -2 \\ -19 \end{bmatrix} + x_3 \begin{bmatrix} -11 \\ 5 \\ 31 \end{bmatrix} = \begin{bmatrix} 6 \\ -3 \\ -17 \end{bmatrix}$

4. $x_1 \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$

♠ 5. yes, $\boldsymbol{b} = 3\boldsymbol{a}_1 + 2\boldsymbol{a}_2$

♠ 6. no, the corresponding linear system is inconsistent

♠ 7. yes, $\boldsymbol{b} = 3\boldsymbol{a}_1 + 4\boldsymbol{a}_2$

8. yes, $\boldsymbol{b} = 3\boldsymbol{a}_1 + \boldsymbol{a}_2$

9. yes, $\boldsymbol{a}_1 + 2\boldsymbol{a}_2 + 3\boldsymbol{a}_3 + 4\boldsymbol{a}_4$

10. no

♠ 11. yes, $\boldsymbol{b} = -\boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3$, for example. There are, in fact, infinitely many possible linear combinations. A dependency relation occurs whenever $x_1 = 2 - x_3$ and $x_2 = 3 - 2x_3$.

12. yes, $\boldsymbol{b} = \boldsymbol{a}_1 + 4\boldsymbol{a}_3$, for example.

♠ 13. Hint: Although we know nothing about the vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, we can still control the scalars $x_i \in F$ such that

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \ldots + x_n\boldsymbol{v}_n = \boldsymbol{0}.$$

What should we set x_i equal to?

14. Hint: To show that $\text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} \subseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ take an arbitrary vector in $\text{Span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$, namely $a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2$ for scalars $a_1, a_2 \in F$, and argue why it is in $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, that is, $a_1\boldsymbol{u}_1 + a_2\boldsymbol{u}_2 = b_1\boldsymbol{v}_1 + b_2\boldsymbol{v}_2 + b_3\boldsymbol{v}_3$ for some scalars $b_1, b_2, b_3 \in F$. By assumption we know that $\boldsymbol{u}_1 = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3$ and $\boldsymbol{u}_2 = d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + d_3\boldsymbol{v}_3$.

15. Hint: Show that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \subseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ and that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \supseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$. That is, take an arbitrary vector in one set and argue why it is in the other set. To show $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \subseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, set one of the coefficients to zero. To show $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} \supseteq \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ then suppose $\boldsymbol{v}_3 = a\boldsymbol{v}_1 + b\boldsymbol{v}_2$ and combine like-terms. Or, if you already did Exercise 14, you could apply that exercise twice for a very fast argument.

16. Hint: Use Exercise 15 to show that $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2\}$ and $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_2\}$.

17. Hint: Generalize Exercise 14.

18. Hint: Generalize Exercise 15.

19. Hint: Generalize Exercise 16.

20. Hint: Use Exercise 19.

2.2 Matrix Equations

1. False. *A homogeneous linear system can be written as $A\boldsymbol{x} = \boldsymbol{0}$, where A is the $m \times n$ coefficient matrix, $\boldsymbol{x} \in F^n$, and $\boldsymbol{0} \in F^m$ is the zero vector.*

2. $\begin{bmatrix} 30 \\ 5 \\ 9 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 11 \\ 6 \\ -2 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 10 + 10i \\ -2 - 11i \end{bmatrix}$

♠ 5. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

♠ 6. $\begin{cases} 4x_1 + x_2 + 2x_3 \equiv 1 \\ 4x_2 + 3x_3 + x_4 \equiv 0 \pmod 5 \\ 2x_1 + 2x_2 + 4x_3 \equiv 2 \end{cases}$

♠ 7. $\begin{cases} x + 2y = 2 \\ 3x + 4y = 0 \\ -x - 2y = 2 \\ 4x + 5y = -9 \end{cases}$

8. $\begin{cases} 21x + 27y + 17z - 32w = 12 \\ 4x + 42y + 6z + 19w = -6 \\ 103x - 72y + 17z - 8w = 15 \\ 2x + 11y - 10z + 13w = 1 \end{cases}$

9. $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$

♠ 10. $\begin{bmatrix} -7 \\ 9 \\ 2 \end{bmatrix}$

♠ 11. inconsistent

12. $\begin{bmatrix} 1+i \\ 3 \\ 1-i \end{bmatrix}$

♠ 13. $\begin{bmatrix} 4 \\ 2 \\ 4 \end{bmatrix}$

14. $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 \\ 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

♠ 17. $T(1,0,0,1) = (5,1)$, $T(1,2,3,4) = (12,10)$

♠ 18. $\boldsymbol{x} = (2/3, 1/3, 0, 0)$ (Answers may vary)

19. The solution would also be $(4,3,2,1)$ because the system of linear equations and the matrix equation represent the same linear system and hence have the same solution set.

2.3 Linear Independence

1. False. *If a set $S \subseteq F^n$ contains the zero vector, then S is linearly dependent.*

2. True. 3. True. 4. True. 5. True. 6. True.

♠ 7. linearly dependent; the set contains **0**.

♠ 8. linearly independent; the second vector is not a scalar multiple of the first.

9. linearly dependent; the second vector is a scalar multiple of the first.

10. linearly dependent; there are more vectors than components (underdetermined system).

♠ 11. linearly independent; it is a single nonzero vector.

♠ 12. linearly dependent; there are more vectors than components.

13. linearly dependent; the second vector is the zero vector.

14. linearly independent

♠ 15. linearly dependent; (Answers may vary)

$$2\begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ -2 \\ 6 \end{bmatrix} + 0\begin{bmatrix} -2 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

♠ 16. linearly independent

17. linearly independent

♠ 18. linearly independent

♠ 19. linearly independent

♠ 20. linearly dependent; (Answers may vary)

$$2\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 6 \\ 6 \\ 4 \end{bmatrix} + 4\begin{bmatrix} 6 \\ 5 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 5 \\ 4 \\ 4 \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \pmod 7$$

2.4 Affine Geometry

♠ 1. F^2 ♠ 2. F ♠ 3. F^8 ♠ 4. F^3 ♠ 5. F^5 ♠ 6. F^0

7. $\boldsymbol{x} = \begin{bmatrix} 5 \\ 4 \\ 6 \end{bmatrix} + t \begin{bmatrix} 4 \\ 6 \\ 9 \end{bmatrix}$,

$$\begin{cases} x_1 = 5 + 4t \\ x_2 = 4 + 6t \\ x_3 = 6 + 9t \end{cases}$$

8. $\boldsymbol{x} = \begin{bmatrix} -3 \\ 6 \\ 1 \end{bmatrix} + a \begin{bmatrix} 1 \\ 6 \\ 5 \end{bmatrix} + b \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}$,

$$\begin{cases} x_1 = -3 + a + 4b \\ x_2 = 6 + 6a - 2b \\ x_3 = 1 + 5a + 3b \end{cases}$$

9. $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + s \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix} + t \begin{bmatrix} -4 \\ 3 \\ -2 \end{bmatrix}$,

$$\begin{cases} x_1 = 1 - s - 4t \\ x_2 = 2 + s + 3t \\ x_3 = 3 - s - 2t \end{cases}$$

♠ 10. $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} + s \begin{bmatrix} 1 \\ 1 \\ 3 \\ 3 \end{bmatrix} + t \begin{bmatrix} 2 \\ 4 \\ 2 \\ 0 \end{bmatrix}$,

$$\begin{cases} x_1 = 1 + s + 2t \\ x_2 = 2 + s + 4t \\ x_3 = 3 + 3s + 2t \\ x_4 = 4 + 3s \end{cases}$$

♠ 11. $\boldsymbol{x} = \begin{bmatrix} 0 \\ 1 \\ 4 \\ -1 \end{bmatrix} + r \begin{bmatrix} 11 \\ -1 \\ -4 \\ 3 \end{bmatrix} + s \begin{bmatrix} -2 \\ -4 \\ -4 \\ 2 \end{bmatrix} + t \begin{bmatrix} 9 \\ 8 \\ -4 \\ 0 \end{bmatrix}$,

$$\begin{cases} x_1 = 11r - 2s + 9t \\ x_2 = 1 - r - 4s + 8t \\ x_3 = 4 - 4r - 4s - 4t \\ x_4 = -1 + 3r + 2s \end{cases}$$

♠ 12. $\boldsymbol{x} = \begin{bmatrix} i \\ 1+2i \end{bmatrix} + t \begin{bmatrix} 2-i \\ 5 \end{bmatrix}$,

$$\begin{cases} x_1 = i + (2-i)t \\ x_2 = 1 + 2i + 5t \end{cases}$$

♠ 13. $\boldsymbol{x} \equiv \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \pmod 5$,

$$\begin{cases} x_1 \equiv 1 + t \\ x_2 \equiv 2 + t \pmod 5 \\ x_3 \equiv 3 \end{cases}$$

♠ 14. $\boldsymbol{x} \equiv \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} + s \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 5 \end{bmatrix} \pmod 7$,

$$\begin{cases} x_1 \equiv 2 + s \\ x_2 \equiv 3 + 2s \pmod 7 \\ x_3 \equiv 1 + 3s + 5t \end{cases}$$

15. $\boldsymbol{x} = \begin{bmatrix} 1 \\ -2 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ -2 \\ 5 \\ 1 \end{bmatrix}$,

$$\begin{cases} x_1 = 1 + 2t \\ x_2 = -2 - 2t \\ x_3 = 4 + 5t \\ x_4 = t \end{cases}$$

♠ 16. $\boldsymbol{x} = \begin{bmatrix} -7 \\ -7 \\ 4 \\ 0 \end{bmatrix} + t \begin{bmatrix} 17 \\ 12 \\ -3 \\ 1 \end{bmatrix}$,

$$\begin{cases} x_1 = -7 + 17t \\ x_2 = -7 + 12t \\ x_3 = 4 - 3t \\ x_4 = t \end{cases}$$

17. $\boldsymbol{x} = \begin{bmatrix} -3 \\ 2 \\ -6 \end{bmatrix}$

♠ 18. $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

19. $\begin{bmatrix} 17 \\ 22 \\ 33 \end{bmatrix}$

2.5 Subspaces

♠ 1. It is not a subspace because $(0,0) \notin W$. Likewise, it is not closed under addition, e.g. $(1,0),(0,1) \in W$ but $(1,0)+(0,1) = (1,1) \notin W$. It is likewise not closed under scalar multiples, e.g. $(1,0) \in W$ but $2(1,0) = (2,0) \notin W$.

♠ 2. It is not a subspace because it is not closed under addition, e.g. $(1,0),(0,1) \in W$ but $(1,0)+(0,1) = (1,1) \notin W$. Of course, there are some instances where the sum is contained, e.g. $(0.1,0.2)+(-0.1,0.3) = (0,0.5) \in W$, but closure under addition means that every possible sum is contained not just some sums. It is likewise not closed under scalar multiples, e.g. $(1,0) \in W$ but $2(1,0) = (2,0) \notin W$. It does contain the zero vector.

♠ 3. It is not a subspace because it is not closed under addition. Note that $(1,1),(2,4) \in W$ but $(1,1)+(2,4) = (3,5) \notin W$. It likewise fails closure under scalar multiples since $(1,1) \in W$ but $2(1,1) = (2,2) \notin W$. On the other hand, it does contain the zero vector.

♠ 4. It is not a subspace because the sum of two points in W need not belong to W, for example, $(1,0)+(0,1) = (1,1) \notin W$. On the other hand, $(0,0) \in W$ and W is closed under scalar multiples.

♠ 5. It is not a subspace because it is not closed under scalar multiples, e.g. $(1,1) \in W$ but $(-1)(1,1) = (-1,-1) \notin W$. Is is closed under addition and contains the zero vector.

6. It is not a subspace because $(0,0) \notin W$.

♠ 7. It is not a subspace because the sum of two such functions will have a y-intercept of 2, not 1; so, the set is not closed under addition. Also, it is not closed under scalar multiples nor contains the zero function.

♠ 8. It is a subspace, if $f(0) = 0$ and $g(0) = 0$ then $(f+g)(0) = f(0)+g(0) = 0+0 = 0$ and $(cf)(0) = c[f(0)] = c0 = 0$. Of course, the zero function is a member.

♠ 9. It is not a subspace, because if the end behavior of f is ∞ then the end behavior of $-f$ will be ∞; so, the set is not closed under scalar multiples. It likewise does not contain the zero function. It is closed under addition though.

♠ 10. It is a subspace, because if $f(-x) = -f(x)$ and $g(-x) = -g(x)$ then $(f+g)(-x) = f(-x)+g(-x) = -f(x)-g(x) = -[f(x)+g(x)] = -(f+g)(x)$ and $(cf)(-x) = c[f(-x)] = c[-f(x)] = -[cf(x)] = -(cf)(x)$. Of course, the zero function is a member.

2.6 Solution Sets of Linear Systems

1. Yes; x_3

2. $\text{Span}\left\{\begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix}\right\}$

♠ 3. $\text{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}\right\}$

4. $\text{Span}\left\{\begin{bmatrix} 2 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 5 \\ -4 \\ 0 \\ 3 \end{bmatrix}\right\}$

♠ 5. $\text{Span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}$

♠ 6. $\text{Span}\left\{\begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 3 \\ -4 \end{bmatrix}\right\}$

♠ 7. $\text{Span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -5 \\ 0 \\ 0 \\ 0 \end{bmatrix}\right\}$

8. Span $\left\{ \begin{bmatrix} -2 \\ 5 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

9. $\boldsymbol{x} = \begin{bmatrix} 5 \\ 0 \\ 0 \\ -8 \end{bmatrix} + s \begin{bmatrix} 3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -4 \\ 0 \\ 1 \\ 0 \end{bmatrix}$

♠ 10. $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$

♠ 11. $\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

♠ 12. $\boldsymbol{x} = \begin{bmatrix} 1+i \\ 3+4i \\ 0 \end{bmatrix} + t \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix}$

♠ 13. $\boldsymbol{x} \equiv \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

♠ 14. We saw in Example 2.6.6 that $\boldsymbol{x}_0 = \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}$ and $\boldsymbol{x}_1 = \begin{bmatrix} 5 \\ 2 \\ 1 \end{bmatrix}$ are two solution to the linear system, but $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_1 \equiv \begin{bmatrix} 4 \\ 4 \\ 1 \end{bmatrix} \pmod 7$, which is not a solution $\boldsymbol{x}$ since

$$A\boldsymbol{x} = \begin{bmatrix} 3 & 5 & 3 \\ 4 & 5 & 4 \\ 6 & 1 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 3(4)+5(4)+3(1) \\ 4(4)+5(4)+4(1) \\ 6(4)+1(4)+6(1) \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 5 \\ 6 \end{bmatrix} \not\equiv \begin{bmatrix} 0 \\ 6 \\ 3 \end{bmatrix}.$$

♠ 15. We saw in Example 2.6.6 that $\boldsymbol{x}_0 = \begin{bmatrix} 6 \\ 2 \\ 0 \end{bmatrix}$. Then $2\boldsymbol{x}_0 \equiv \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix} \pmod 7$. But then

$$A\boldsymbol{x} = \begin{bmatrix} 3 & 5 & 3 \\ 4 & 5 & 4 \\ 6 & 1 & 6 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 3(5)+5(4)+3(0) \\ 4(5)+5(4)+4(0) \\ 6(5)+1(4)+6(0) \end{bmatrix} \equiv \begin{bmatrix} 0 \\ 5 \\ 6 \end{bmatrix} \not\equiv \begin{bmatrix} 0 \\ 6 \\ 3 \end{bmatrix}.$$

♠ 16. Hint: Suppose $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are solutions to the linear system $A\boldsymbol{x} = \boldsymbol{b}$. Multiply the equation $\boldsymbol{x} = (1-t)\boldsymbol{x}_0 + t\boldsymbol{x}_1$ by the matrix A, using also the fact that $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are solutions to the linear system. Recall that multiplication by a matrix distributes across vector addition and commutes with scalar multiplication. It is a linear transformation, after all.

17. Hint: Like the hint to the previous exercise, suppose $\boldsymbol{x} = a_0\boldsymbol{x}_0 + a_1\boldsymbol{x}_1 + \ldots + a_m\boldsymbol{x}_m$ such that $\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ are solutions to $A\boldsymbol{x} = \boldsymbol{b}$ and $a_0 + a_1 \ldots + a_m = 1$. Then simplify

$$A(a_0\boldsymbol{x}_0 + a_1\boldsymbol{x}_1 + \ldots + a_m\boldsymbol{x}_m).$$

2.7 Bases

1. False. *If A is an $m \times n$ matrix, then* $\operatorname{rank}(A) + \operatorname{nullity}(A) = \boldsymbol{n}$.

2. $W = \operatorname{Span}\left\{\begin{bmatrix}1\\2\\7\end{bmatrix}, \begin{bmatrix}21\\16\\43\end{bmatrix}\right\}$

3. $W = \operatorname{Span}\left\{\begin{bmatrix}0\\0\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right\}$

4. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 0$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}15\\9\\0\end{bmatrix}, \begin{bmatrix}8\\6\\2\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\{\}$

5. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 0$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}1\\2\\4\end{bmatrix}, \begin{bmatrix}1\\3\\2\end{bmatrix}, \begin{bmatrix}2\\5\\4\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\{\}$

6. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 1$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}1\\3\\1\end{bmatrix}, \begin{bmatrix}1\\4\\3\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}3\\2\\-5\\4\end{bmatrix}\right\}$

7. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 1$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}5\\4\\2\end{bmatrix}, \begin{bmatrix}4\\4\\5\end{bmatrix}, \begin{bmatrix}2\\2\\6\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}7\\-13\\8\\1\end{bmatrix}\right\}$

♠ 8. $\operatorname{rank} A = 2$, $\operatorname{nullity} A = 3$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}8\\-2\end{bmatrix}, \begin{bmatrix}-3\\1\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}2\\1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}-3\\-3\\0\\1\\0\end{bmatrix}, \begin{bmatrix}-4\\-5\\0\\0\\1\end{bmatrix}\right\}$

♠ 9. $\operatorname{rank} A = 2$, $\operatorname{nullity} A = 2$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}-1\\7\\-3\end{bmatrix}, \begin{bmatrix}0\\-8\\2\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\\0\end{bmatrix}\right\}$

10. $\operatorname{rank} A = 2$, $\operatorname{nullity} A = 2$,
$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}\right\}$,
$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}3\\-4\\1\\0\end{bmatrix}, \begin{bmatrix}17\\-13\\0\\1\end{bmatrix}\right\}$

11. rank $A = 4$, nullity $A = 2$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}6\\5\\2\\0\end{bmatrix}, \begin{bmatrix}5\\4\\4\\1\end{bmatrix}, \begin{bmatrix}4\\3\\6\\0\end{bmatrix}, \begin{bmatrix}2\\1\\2\\2\end{bmatrix}\right\}, \quad \text{Nul}(A) = \text{Span}\left\{\begin{bmatrix}1\\-1\\-1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}23\\-50\\18\\0\\19\\2\end{bmatrix}\right\}$$

12. rank $A = 2$, nullity $A = 3$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1\\2\\3\\2\end{bmatrix}, \begin{bmatrix}2\\4\\6\\-2\end{bmatrix}\right\},$$

$$\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix}7\\1\\3\\0\\0\end{bmatrix}, \begin{bmatrix}-8\\-2\\0\\3\\0\end{bmatrix}, \begin{bmatrix}13\\1\\0\\0\\3\end{bmatrix}\right\}$$

13. rank $A = 2$, nullity $= 0$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1+i\\2+4i\\6-10i\\-8-2i\end{bmatrix}, \begin{bmatrix}4-i\\4-2i\\3+i\\12+10i\end{bmatrix}\right\},$$

$$\text{Nul}(A) = \text{Span}\{\}$$

♠ 14. rank $A = 3$, nullity $A = 3$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1+2i\\2+4i\\3-i\end{bmatrix}, \begin{bmatrix}-1\\1+4i\\2\end{bmatrix}, \begin{bmatrix}0\\-2\\i\end{bmatrix}\right\},$$

$$\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix}i\\1\\0\\0\\0\\0\end{bmatrix}, \begin{bmatrix}-1\\0\\-i\\2\\1\\0\end{bmatrix}, \begin{bmatrix}-1-i\\0\\-2\\-1-2i\\0\\1\end{bmatrix}\right\}$$

♠ 15. rank $A = 3$, nullity $A = 3$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\},$$

$$\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix}1\\1\\1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\\0\\0\\1\end{bmatrix}\right\}$$

16. rank $A = 3$, nullity $A = 3$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}\right\},$$

$$\text{Nul}(A) = \text{Span}\left\{\begin{bmatrix}1\\0\\1\\1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\1\end{bmatrix}\right\}$$

♠ 17. $\operatorname{rank} A = 2$, $\operatorname{nullity} A = 2$,

$$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}2\\4\\2\\0\end{bmatrix}, \begin{bmatrix}1\\0\\3\\1\end{bmatrix}\right\}, \qquad \operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}3\\0\\3\\1\end{bmatrix}\right\}$$

18. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 1$,

$$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}6\\1\\3\end{bmatrix}, \begin{bmatrix}2\\5\\6\end{bmatrix}, \begin{bmatrix}0\\6\\2\end{bmatrix}\right\}, \qquad \operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}0\\-1\\2\\0\end{bmatrix}\right\}$$

19. $\operatorname{rank} A = 3$, $\operatorname{nullity} A = 3$,

$$\operatorname{Col}(A) = \operatorname{Span}\left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}2\\3\\4\end{bmatrix}, \begin{bmatrix}6\\0\\2\end{bmatrix}\right\},$$

$$\operatorname{Nul}(A) = \operatorname{Span}\left\{\begin{bmatrix}12\\-4\\-2\\1\\0\\0\end{bmatrix}, \begin{bmatrix}250\\-86\\-25\\0\\8\\0\end{bmatrix}, \begin{bmatrix}13\\-5\\-1\\0\\0\\1\end{bmatrix}\right\}$$

20. **Hint**: If $\boldsymbol{v}$ is a strongly positive vector but $\boldsymbol{u}$ is not, is there a positive scalar $c \in \mathbb{R}$ that guarantees that $\boldsymbol{u} + c\boldsymbol{v}$ is also strongly positive? How does the span of $\boldsymbol{u}$ and $\boldsymbol{u} + c\boldsymbol{v}$ compare to the span of $\boldsymbol{u}$ and $\boldsymbol{v}$?

2.8 Coordinates

♠ 1. $\begin{bmatrix}-5\\3\end{bmatrix}$ ♠ 2. $\begin{bmatrix}2\\3\\-2\end{bmatrix}$ ♠ 3. $\begin{bmatrix}i\\-2\end{bmatrix}$ ♠ 4. $\begin{bmatrix}1\\9\\11\end{bmatrix}$ ♠ 5. $\begin{bmatrix}-4\\9\\32\\5\end{bmatrix}$ ♠ 6. $\begin{bmatrix}4\\8\\7\end{bmatrix}$

7. $\begin{bmatrix}23+13i\\10+6i\\-5-13i\end{bmatrix}$ ♠ 8. $\begin{bmatrix}3&2&1\\-1&1&1\\0&-2&-4\end{bmatrix}$ ♠ 9. $\begin{bmatrix}3&2&3&7\\10&4&0&-1\\-5&6&-4&-2\\0&8&5&7\end{bmatrix}$ 10. $\begin{bmatrix}1&-2&6&-2\\-17&-7&-8&10\\69&47&26&-20\\-17&-11&-8&6\end{bmatrix}$

♠ 11. $\begin{bmatrix}1+i&3\\i&2+2i\end{bmatrix}$ 12. $\begin{bmatrix}0&3&1\\3&1&4\\2&4&2\end{bmatrix}$ 13. $\begin{bmatrix}0&3&2\\4&4&0\\2&4&1\end{bmatrix}$

Chapter 3 : The Algebra and Geometry of Matrices

3.1 Matrix Operations

1. True

2. False. *If A is a square matrix, then the trace of A is the sum of the* ***diagonal*** *entries of A.*

3. True 4. True 5. True 6. True 7. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

8. There are more columns in C then rows in A.

9. There are fewer columns in AB then rows in D.

10. C is not square.

♠ 11. $\begin{bmatrix} 2 & -1 & 8 \\ 13 & -11 & 0 \\ 8 & -2 & -5 \end{bmatrix}$

♠ 12. $\begin{bmatrix} 4 & -7 & -6 \\ 31 & 3 & -6 \\ -2 & -8 & 5 \end{bmatrix}$

♠ 13. $\begin{bmatrix} 14 & 23 & 21 & -16 \\ 9 & -21 & -28 & 17 \\ -2 & -17 & -19 & 12 \end{bmatrix}$

♠ 14. $\begin{bmatrix} 1 & 4 & 1 \\ 2 & -3 & -2 \\ 3 & 0 & -1 \end{bmatrix}$

♠ 15. $\begin{bmatrix} 3 & 1 & 3 \\ 0 & 7 & 3 \\ -1 & 8 & 2 \\ 2 & -3 & -4 \end{bmatrix}$

♠ 16. $\begin{bmatrix} 2 & -5 & 0 & 1 \\ 3 & 0 & -2 & 2 \\ -2 & 7 & 4 & 3 \end{bmatrix}$

♠ 17. -3

♠ 18. -8

♠ 19. $\begin{bmatrix} 3 & -29 & 10 \\ 42 & 39 & -31 \\ 14 & 18 & -12 \\ 28 & -9 & -7 \end{bmatrix}$

♠ 20. $\begin{bmatrix} 9 & -6 & 6 \\ 0 & 6 & 12 \\ -6 & 6 & -3 \end{bmatrix}$

21. $\begin{bmatrix} 3 & 3 \\ 2 & 8 \\ 7 & 14 \end{bmatrix}$

22. $\begin{bmatrix} 1 & 7 \\ 26 & 27 \end{bmatrix}$

23. $\begin{bmatrix} -7 & -2 \\ 2 & -4 \\ 6 & -8 \end{bmatrix}$

24. $\begin{bmatrix} 19 & -14 \\ 29 & -54 \end{bmatrix}$

♠ 25. $\begin{bmatrix} -1+3i & -5+10i \\ 5-4i & -1-3i \end{bmatrix}$

♠ 26. $\begin{bmatrix} 3+6i & 3-9i \\ -2+11i & 5+3i \\ 6+22i & 10-6i \end{bmatrix}$

♠ 27. $\begin{bmatrix} 1-2i & 5 \\ 1+3i & 1+i \end{bmatrix}$

♠ 28. $\begin{bmatrix} 3 & -i & 3-i \\ 0 & -2i & 1-3i \end{bmatrix}$

♠ 29. $\begin{bmatrix} 1 & 5 \\ 0 & 4-i \\ -i & 0 \\ 1+2i & -4 \end{bmatrix}$

♠ 30. $2+i$

♠ 31. $1+i$

♠ 32. $\begin{bmatrix} 15+20i & 5-8i \\ 16+13i & 3-5i \\ 0 & -3 \\ -12-16i & 2+i \end{bmatrix}$

♠ 33. $\begin{bmatrix} 3-11i & 5-5i \\ 10+5i & 6-17i \end{bmatrix}$

♠ 34. $\begin{bmatrix} 2 & 1 & 3 & 2 \\ 4 & 3 & 0 & 1 \\ 4 & 3 & 0 & 1 \\ 1 & 4 & 4 & 3 \end{bmatrix}$

♠ 35. $\begin{bmatrix} 0 & 2 & 3 & 0 \\ 2 & 0 & 4 & 2 \\ 4 & 0 & 2 & 0 \end{bmatrix}$

♠ 36. $\begin{bmatrix} 1 & 0 & 1 & 4 \\ 2 & 2 & 2 & 4 \\ 3 & 2 & 2 & 3 \\ 4 & 3 & 1 & 2 \end{bmatrix}$

♠ 37. $\begin{bmatrix} 3 & 1 & 3 \\ 0 & 2 & 3 \\ 4 & 3 & 2 \\ 2 & 2 & 1 \end{bmatrix}$

♠ 38. $\begin{bmatrix} 2 & 0 & 0 & 1 \\ 3 & 0 & 3 & 2 \\ 3 & 2 & 4 & 3 \end{bmatrix}$

♠ 39. 2

♠ 40. 2

♠ 41. $\begin{bmatrix} 4 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 4 & 3 & 0 \end{bmatrix}$

♠ 42. $\begin{bmatrix} 0 & 0 & 3 & 0 \\ 4 & 1 & 4 & 2 \\ 3 & 1 & 0 & 0 \\ 4 & 4 & 0 & 1 \end{bmatrix}$

43. $7, \begin{bmatrix} 8 & 0 & 3 \\ -7 & -2 & 2 \\ 2 & 4 & 1 \end{bmatrix}, 7$

44. $2, \begin{bmatrix} 1/8 & 9/8 & 1 & 1/4 \\ -7/8 & -1/2 & -12/15 & 8/15 \\ 2/5 & 1/4 & 2 & -3/8 \\ -12/13 & 1/4 & 1/5 & 3/8 \end{bmatrix}, 2$

45. $6, \begin{bmatrix} 5 & 3 & 5 & 2 \\ 1 & 0 & 2 & 3 \\ 6 & 3 & 1 & 0 \\ 2 & 4 & 1 & 0 \end{bmatrix} \pmod 7, 6$

46. $3+9i, \begin{bmatrix} 1+3i & 2+4i & -1-2i \\ -5+5i & -8i & 3+6i \\ 2-2i & 6-i & 2-4i \end{bmatrix}, 3-9i$

3.2 Matrix Properties

1. For all $A, B, C \in F^{m\times n}$ and all scalars $c, d \in F$, the following eight properties hold:

   (i) $A + B = B + A$

   (ii) $(A + B) + C = A + (B + C)$

   (iii) There exists a matrix $0 \in F^{m\times n}$, such that $A + 0 = 0 + A = A$

   (iv) For each A, there exists a matrix $-A \in F^{m\times n}$, such that $A+(-A) = (-A)+A = 0$

   (v) $c(A + B) = cA + cB$

   (vi) $(c + d)A = cA + dA$

   (vii) $c(dA) = (cd)A$

   (viii) $1A = A$

2. $AB = \begin{bmatrix} 5 & 12 & 21 \\ 1 & 8 & 22 \\ 8 & 15 & 26 \end{bmatrix} \neq \begin{bmatrix} 14 & 19 & 5 \\ 17 & 9 & 8 \\ 35 & 16 & 16 \end{bmatrix} = BA$

♠ 3. $\text{Row}(A) = \text{Span}\{[1\ 0\ -2\ 3\ 4], [0\ 1\ -1\ 3\ 5]\}$, $\text{corank}(A) = 2$

♠ 4. $\text{Row}(A) = \text{Span}\{[1\ -2\ -1\ 0], [0\ 0\ 0\ 1]\}$, $\text{corank}(A) = 2$

♠ 5. $\text{Row}(A) = \text{Span}\{[1\ 0\ 0\ 1\ 1\ 1], [0\ 1\ 0\ 1\ 1\ 0], [0\ 0\ 1\ 1\ 0\ 1]\}$, $\text{corank}(A) = 3$

♠ 6. $\text{Row}(A) = \text{Span}\{[1\ 3\ 0\ 2], [0\ 0\ 1\ 2]\}$, $\text{corank}(A) = 2$

♠ 7. $\text{Row}(A) = \text{Span}\{[1\ i\ 0\ 0\ 1\ 1-i], [0\ 0\ 1\ 0\ -i\ 2], [0\ 0\ 0\ 1\ -2\ 1-2i]\}$, $\text{corank}(A) = 3$

8. $\text{rank}(A) = \text{corank}(A) = 3$,

$$\text{Col}(A) = \text{Span}\left\{\begin{bmatrix}1\\1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\2\\-1\\-3\\-2\end{bmatrix}, \begin{bmatrix}7\\10\\1\\5\\0\end{bmatrix}\right\},$$

$\text{Row}(A) = \text{Span}\{(1,1,-3,7,9,-9), (0,1,-4,3,4,-3), (0,0,0,1,-1,2)\}$

9. Hint: Is matrix multiplication commutative? The multiplication of a field should be commutative. Can you find an example of two 2×2 matrices that do not commute? I hope so, because the section provides some.

10. Hint: If $n = 1$, then $F^{1\times 1}$ would be the set of 1×1 matrices, which are really just scalars. Hence, $F^{1\times 1}$ is really just F itself, which is a field by assumption. If $n > 1$, then consider the following two matrices:

$$\begin{bmatrix}1&0&0&0&\cdots\\0&0&0&0&\cdots\\0&0&0&0&\cdots\\0&0&0&0&\cdots\\\vdots&\vdots&\vdots&\vdots&\ddots\end{bmatrix}, \begin{bmatrix}0&1&0&0&\cdots\\0&0&0&0&\cdots\\0&0&0&0&\cdots\\0&0&0&0&\cdots\\\vdots&\vdots&\vdots&\vdots&\ddots\end{bmatrix}.$$

3.3 Matrix Inverses

1. $\begin{bmatrix}3&-5\\-1&2\end{bmatrix}$

2. $\frac{1}{2}\begin{bmatrix}-2&1\\3&-1\end{bmatrix}$

3. $\frac{1}{17}\begin{bmatrix}5&4\\7&9\end{bmatrix}$

♠ 4. $\begin{bmatrix}3&-5\\-1&2\end{bmatrix}$

♠ 5. $\begin{bmatrix}-7&5\\3&-2\end{bmatrix}$

♠ 6. $\frac{1}{10}\begin{bmatrix}1&-3\\2&4\end{bmatrix}$

♠ 7. $\begin{bmatrix}4&1\\2&1\end{bmatrix}$

♠ 8. $\begin{bmatrix}4&1\\8&6\end{bmatrix}$

♠ 9. $\frac{1}{1+i}\begin{bmatrix}2&-1-i\\i&1\end{bmatrix}$

10. $\frac{1}{2}\begin{bmatrix}-3\\2\end{bmatrix}$

♠ 11. $\begin{bmatrix}29\\-25\\7\end{bmatrix}$

♠ 12. $\begin{bmatrix}18\\15\\-11\\8\end{bmatrix}$

♠ 13. $(0,1,1)$

♠ 14. nonsinuglar, the matrix has full rank

♠ 15. singular, it is not a square matrix

♠ 16. nonsingular, the mapping $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is surjective (onto)

♠ 17. singular, A is not row equivalent to I_5

♠ 18. not enough information, if the RREF is I_3 it is nonsingular, otherwise it is singular.

♠ 19. nonsingular, it has 3 pivot columns

♠ 20. singular, the mapping $\boldsymbol{x} \mapsto A\boldsymbol{x}$ is not injective (one-to-one)

♠ 21. singular, A is not row equivalent to I_n, otherwise this would suppose that I_n is row equivalent to a singular matrix (a contradiction)

♠ 22. singular, the columns of A are linearly dependent

23. $X = (CBA)^{-1}$

♠ 24. $X = -A^{-1}B$

♠ 25. $X = BA - A^{-1}C$

♠ 26. $X = BC^{-1}A^{-1}$

♠ 27. $X = C^{-1}B - A$

28. $X = D^{-1}(ECA + B)$

29. *Proof of Corollary 3.3.8.* By Theorem 3.3.7, $A^\top$ is invertible and products of invertible matrices are invertible. Therefore, $AA^\top$ and $A^\top A$ are invertible. ∎

30. *Proof of Corollary 3.3.12.* If AB is invertible, then there exists some matrix C such that $(AB)C = I_n$. Thus, $A(BC) = I_n$, which shows that A is invertible by (iv) in the Nonsingular Matrix Theorem. Similarly, $C(AB) = I_n$, which implies that $(CA)B = I_n$. Therefore, B is likewise invertible by (iii) in the Nonsingular Matrix Theorem. ∎

3.4 Elementary Matrices

♠ 1. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

♠ 6. $\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

7. singular

8. $\begin{bmatrix} -2 & 7 & -1 \\ -2 & 6 & -1 \\ -3 & 10 & -1 \end{bmatrix}$

♠ 9. $\frac{1}{6}\begin{bmatrix} 3 & 3 & -3 \\ -6 & -12 & 24 \\ 5 & 7 & -9 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 2 & 3 & 2 \\ 1 & 4 & 2 \\ 1 & 1 & 1 \end{bmatrix}$

11. $\begin{bmatrix} -\frac{25}{8} & 4 & -\frac{3}{8} \\ -2 & 2 & 0 \\ -\frac{3}{8} & 0 & -\frac{1}{8} \end{bmatrix}$

12. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

♠ 13. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -6 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

♠ 14. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 12 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -16 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2i & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4i \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3i \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

16. $X \equiv \begin{bmatrix} 3 & 6 & 2 \\ 1 & 1 & 5 \\ 3 & 3 & 6 \end{bmatrix} \pmod 7$

3.5 Matrix Factorizations

1. scalar
2. not scalar
3. not scalar
4. scalar

♠ 5. $\begin{bmatrix} 2 & 4 \\ 3 & -2 \\ -3 & -15 \end{bmatrix}$

♠ 6. $\begin{bmatrix} -9 & 6 & 18 \\ 0 & 5 & -50 \\ 12 & 4 & 12 \end{bmatrix}$

♠ 7. $\begin{bmatrix} 0 & 6 & 0 \\ 0 & 0 & 0 \\ 0 & 15 & -25 \end{bmatrix}$

♠ 8. $A^2 = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 25 \end{bmatrix}$, $A^3 = \begin{bmatrix} 8 & 0 & 0 \\ 0 & -27 & 0 \\ 0 & 0 & 125 \end{bmatrix}$, $A^{-1} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & 1/5 \end{bmatrix}$, $A^{-3} = \begin{bmatrix} 1/8 & 0 & 0 \\ 0 & -1/27 & 0 \\ 0 & 0 & 1/125 \end{bmatrix}$

9. $P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$;

$PA = \begin{bmatrix} 8 & 6 & 3 & 7 \\ 1 & 5 & 8 & 10 \\ 10 & 9 & 7 & 4 \\ 5 & 2 & 6 & 9 \end{bmatrix}$,
Row 1 goes to Row 2,
Row 2 goes to Row 4,
Row 3 goes to Row 1,
Row 4 goes to Row 3;

$AP = \begin{bmatrix} 5 & 10 & 1 & 8 \\ 2 & 9 & 5 & 6 \\ 6 & 7 & 8 & 3 \\ 9 & 4 & 10 & 7 \end{bmatrix}$,
Column 1 goes to Column 3,
Column 2 goes to Column 1,
Column 3 goes to Column 4,
Column 4 goes to Column 2.

♠ 10. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix}$

♠ 11. $\begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 9 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 3 & \frac{2}{5} & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -1 \\ 0 & -10 & 8 \\ 0 & 0 & \frac{49}{5} \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 1 & -3 & 1 & 0 \\ -3 & 4 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$

♠ 14. $LU = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 2 & -5 \\ 0 & 3 \end{bmatrix}$, $\left[\begin{array}{c|c} L & \boldsymbol{b} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & -6 \end{array}\right]$, $\left[\begin{array}{c|c} U & \boldsymbol{y} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & -6 \\ 0 & 1 & -2 \end{array}\right]$

♠ 15. $LU = \begin{bmatrix} 1 & 0 \\ -\frac{3}{2} & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 \\ 0 & 3 \end{bmatrix}$, $\left[\begin{array}{c|c} L & \boldsymbol{b} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 4 \\ 0 & 1 & 15 \end{array}\right]$, $\left[\begin{array}{c|c} U & \boldsymbol{y} \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 7 \\ 0 & 1 & 5 \end{array}\right]$

16. $LU = \begin{bmatrix} 1 & 0 & 0 \\ 8 & 2 & 0 \\ 6 & 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 6 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $\left[\begin{array}{c|c} L & \boldsymbol{b} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -51 \\ 0 & 0 & 1 & 81 \end{array}\right]$, $\left[\begin{array}{c|c} U & \boldsymbol{y} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 54 \\ 0 & 1 & 0 & -17 \\ 0 & 0 & 1 & 27 \end{array}\right]$

♠ 17. $LU = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1/2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$, $\left[\begin{array}{c|c} L & \boldsymbol{b} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & -10 \\ 0 & 0 & 1 & -6 \end{array}\right]$, $\left[\begin{array}{c|c} U & \boldsymbol{y} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 6 \\ 0 & 1 & 0 & -10 \\ 0 & 0 & 1 & -2 \end{array}\right]$

♠ 18. $LU = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{bmatrix}$, $\left[\begin{array}{c|c} L & \boldsymbol{b} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{array}\right]$, $\left[\begin{array}{c|c} U & \boldsymbol{y} \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 6 \end{array}\right]$

19. *Proof.* Let $A = [a_{ij}]$. Since A is lower triangular we have that $a_{ij} = 0$ for $i > j$. On the other hand, say that $A^\top = [b_{ij}]$. Now, $A^\top$ is an upper triangular matrix if $b_{ij} = 0$ whenever $i < j$. By the transpose operation, we have that $b_{ij} = a_{ji}$. Suppose that $i < j$ (or $j > i$). Then we have $a_{ji} = 0$. Thus, $b_{ij} = 0$, that is, $A^\top$ is upper triangular. ∎

3.6 Linear Transformations on $\mathbb{R}^2$

♠ 1. $\begin{bmatrix} 1/2 & 0 \\ 0 & 5 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1 & 0 \\ 2 & 5 \end{bmatrix}$

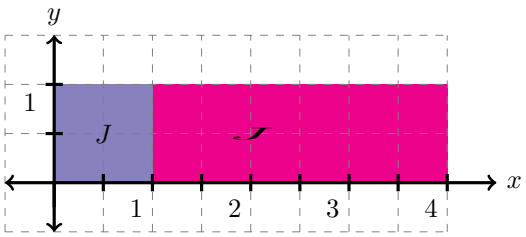
♠ 3. $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 0 & 1 \\ -5 & 0 \end{bmatrix}$

♠ 5. $\frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 \\ 3-6\sqrt{3} & 6+3\sqrt{3} \end{bmatrix}$

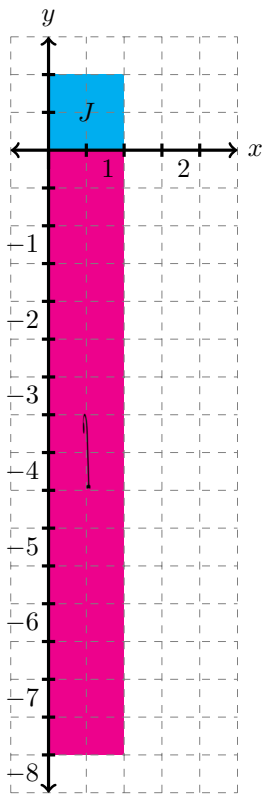
♠ 6. $\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$

Stretch horizontally by a factor of 4.



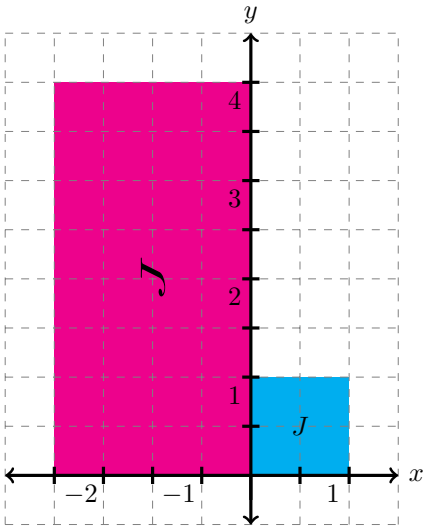
♠ 7. $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 8 \end{bmatrix}$

Stretch vertically by a factor of 8 and reflect across the x-axis.



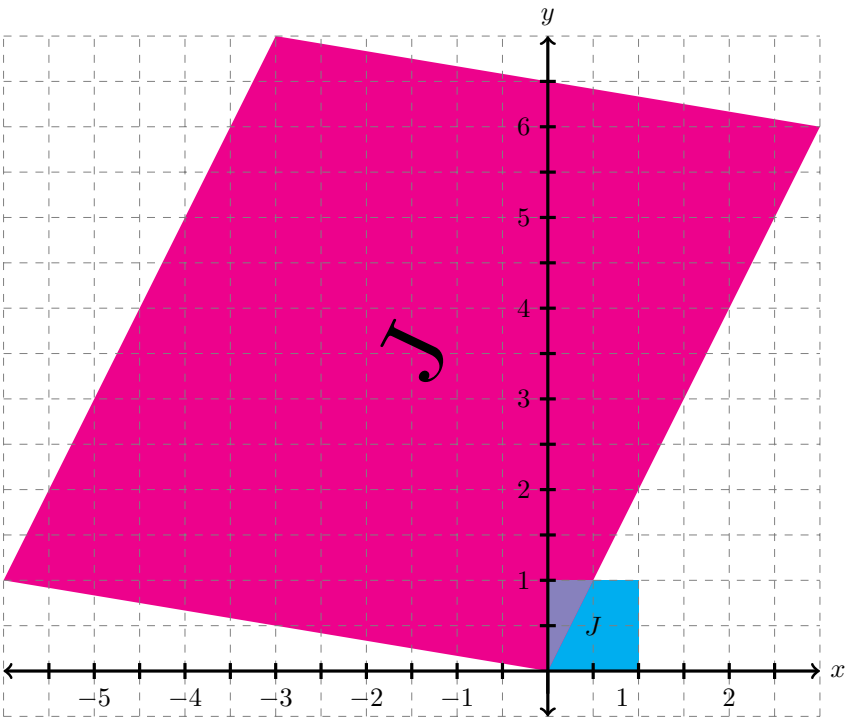
♠ 8. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

Reflect across the x-axis, stretch vertically by a factor of 2, stretch horizontally by a factor of 4, and reflect across the line $y = x$.



♠ 9. $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 13 \end{bmatrix}\begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$

Shear horizontally by a factor of -2, stretch vertically by a factor of 13, shear vertically by a factor of 6, and stretch horizontally by a factor of 3.



10. $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

♠ 11. (a)

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = \begin{bmatrix} \cos^2\theta - \sin^2\theta & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & \sin^2\theta - \cos^2\theta \end{bmatrix} = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} = A$$

(b) The sequence of geometric transformations in the factorization of A rotates the plane clockwise about the origin by an angle of θ, then reflects the plane across the x-axis, then rotates the plane counter-clockwise about the origin by an angle of θ.

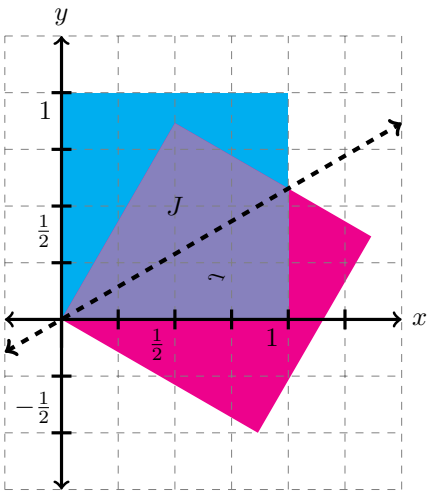
Since the line ℓ is the unique line through the origin that forms an angle of θ with the x-axis, the composite of these transformations moves ℓ onto the x-axis, reflects across the x-axis, and then moves the x-axis back to ℓ. The net effort on the plane is to reflect across the line ℓ.

(c) $\begin{bmatrix} \cos 2(0) & \sin 2(0) \\ \sin 2(0) & -\cos 2(0) \end{bmatrix} = \begin{bmatrix} \cos 0 & \sin 0 \\ \sin 0 & -\cos 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, which is reflection across the x-axis, the unique line which forms a 0 angle with the x-axis.

$$\begin{bmatrix} \cos 2(\pi/4) & \sin 2(\pi/4) \\ \sin 2(\pi/4) & -\cos 2(\pi/4) \end{bmatrix} = \begin{bmatrix} \cos(\pi/2) & \sin(\pi/2) \\ \sin(\pi/2) & -\cos(\pi/2) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$ which is reflection across the line $y = x$, the unique line which forms a $\frac{\pi}{4}$ angle with the x-axis.

$$\begin{bmatrix} \cos 2(\pi/2) & \sin 2(\pi/2) \\ \sin 2(\pi/2) & -\cos 2(\pi/2) \end{bmatrix} = \begin{bmatrix} \cos \pi & \sin \pi \\ \sin \pi & -\cos \pi \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$ which is reflection across the y-axis, the unique line which forms a $\frac{\pi}{2}$ angle with the x-axis.

(d) $$\begin{bmatrix} \cos 2(\pi/6) & \sin 2(\pi/6) \\ \sin 2(\pi/6) & -\cos 2(\pi/6) \end{bmatrix} = \begin{bmatrix} \cos(\pi/3) & \sin(\pi/3) \\ \sin(\pi/3) & -\cos(\pi/3) \end{bmatrix} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & -1/2 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix}$$



(e)

$$B = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}^{-1} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta - m\sin\theta & \sin\theta + m\cos\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta - m\sin\theta\cos\theta + \sin^2\theta & \sin\theta\cos\theta + m\cos^2\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - m\sin^2\theta - \sin\theta\cos\theta & \sin^2\theta + m\sin\theta\cos\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 - \frac{m}{2}\sin 2\theta & m\cos^2\theta \\ -m\sin^2\theta & 1 + \frac{m}{2}\sin 2\theta \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} 2 - m\sin 2\theta & m + m\cos 2\theta \\ -m + m\cos 2\theta & 2 + m\sin 2\theta \end{bmatrix}$$

12. *Proof.* Since A is nonsingular, it can be factored into a product of elementary matrices. Replacement elementary matrices cause shears, scaling elementary matrices cause stretches, compressions, and reflections. Interchange elementary matrices cause reflections. ∎

3.7 Representations of Linear Transformations as Matrices

♠ 1. $[T] = \begin{bmatrix} 2 & -1 \\ 0 & -4 \end{bmatrix}$

♠ 2. $[T] = \begin{bmatrix} 1 & 2 & 1 \\ 3 & -5 & 14 \\ -5 & 3 & -18 \\ -7 & 19 & -40 \end{bmatrix}$

♠ 3. $[T] = \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}$

♠ 4. $[T] = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}$

♠ 5. $[T] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

♠ 6. $[S \circ T] = [S][T] = \begin{bmatrix} 0 & -2 & 1 \\ 1 & 0 & -2 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 3 & 0 \\ 2 & -3 \end{bmatrix} = \begin{bmatrix} -4 & -3 \\ -4 & 8 \end{bmatrix}$

♠ 7. $[T] \sim \begin{bmatrix} \boxed{1} & 0 \\ 0 & \boxed{1} \end{bmatrix}$. Hence, T is one-to-one and onto. $\ker T = \text{Span}\{\} = \{\mathbf{0}\}$, $\text{im}\, T = \text{Span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \mathbb{R}^2$

♠ 8. $[T] \sim \begin{bmatrix} \boxed{1} & 0 & 3 \\ 0 & \boxed{1} & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ Hence, T is neither one-to-one nor onto. $\ker T = \text{Span}\{[-3, 1, 1]^\top\}$, $\text{im}\, T = \text{Span}\{[1, 3, -5, -7]^\top, [2, -5, 3, 19]^\top\}$

♠ 9. $[T] \sim \begin{bmatrix} \boxed{1} \\ 0 \\ 0 \\ 0 \end{bmatrix}$. Hence, T is one-to-one but not onto. $\ker T = \text{Span}\{\} = \{\mathbf{0}\}$, $\text{im}\, T = \text{Span}\{[1, i, -1, -i]^\top\}$

♠ 10. $[T] \sim \begin{bmatrix} \boxed{1} & 1 & 1 & 1 \end{bmatrix}$. Hence, T is onto but not one-to-one. $\ker T = \{[-1, 1, 0, 0]^\top, [-1, 0, 1, 0]^\top, [-1, 0, 0, 1]^\top\}$, $\text{im}\, T = \text{Span}\{1\} = \mathbb{Z}_2$

♠ 11. $[T] \sim \begin{bmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{bmatrix}$. Hence, T is one-to-one and onto. $\ker T = \text{Span}\{\} = \{\mathbf{0}\}$, $\text{im}\, T = \text{Span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \mathbb{Z}_5^3$

12. As this is the identity transformation, it is clearly one-to-one and onto.

13. T cannot be one-to-one as the standard matrix $[T]$ is 2×3 and has too many columns. (It is onto since in the standard matrix, the two rows are not multiples of each other, that is, they are linear independent, and the matrix will have two pivots).

14. T cannot be onto as the standard matrix $[T]$ is 4×3 and has too many rows. (It is onto, something we see from row reducing the standard matrix).

15. T cannot be onto as the standard matrix $[T]$ is 3×1 and has too many rows. (It is one-to-one since only zero maps to the zero vector).

16. T cannot be onto as the standard matrix $[T]$ is 3×1 and has too many rows. (It is one-to-one since only zero maps to the zero vector).

17. T cannot be one-to-one as the standard matrix $[T]$ is 1×3 and has too many columns. (It is onto since for any scalar c we have that $T(c, 0, 0) = c$).

18. As the standard matrix $[T]$ is 2×2, T could potentially be one-to-one or potentially onto. (It is, in fact, both).

Chapter 4 : Orthogonality

4.1 Inner Products

1. 5
3. 18
5. -40
♠ 7. 36
♠ 2. 48
♠ 4. -4
♠ 6. 7
♠ 8. 22
9. 342
♠ 10. -6
♠ 11. $-i$
♠ 12. $1 + 7i$
♠ 13. $-2 + 9i$
♠ 14. $9 - 3i$
15. $26 + 4i$
16. $2 - 13i$
17. $\sqrt{35}$
♠ 18. $5\sqrt{3}$
♠ 19. $3\sqrt{2}$
20. $\sqrt{298}$
21. $\sqrt{17}$
♠ 22. $\sqrt{6}$
23. $\sqrt{14}$
24. $\sqrt{15}$
♠ 25. $\sqrt{22}$
26. $\sqrt{66}$
♠ 27. $3\sqrt{2}$
♠ 28. $3\sqrt{5}$
29. $\sqrt{14}$
30. $\frac{1}{2}$
31. $\pm\frac{\sqrt{23}}{6}$

4.2 Orthogonality

♠ 1. $-\frac{3}{4}$
2. orthogonal
3. no, but $(6, 1) \perp (8, -24)$
4. orthogonal
5. no, but $(2, 8) \perp (4, -1)$
♠ 6. $3x_1 - 2x_2 = -1$
♠ 7. $x_1 - 2x_2 + 2x_3 - x_4 = 0$
♠ 8. $3x_1+6x_2-3x_3-2x_4+5x_5 = 23$
9. $x_1+3x_2+5x_3+7x_4+9x_5 = 6$
10. $\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$
♠ 11. $\left\{\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}\right\}$
♠ 12. $\left\{\begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right\}$
♠ 13. $\left\{\begin{bmatrix} 4 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \end{bmatrix}\right\}$
♠ 14. $\left\{\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right\}$
15. $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$
♠ 16. $\left\{\begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}\right\}$
♠ 17. $\left\{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right\}$
18. $x_1 + 2x_2 + x_3 = 16$

4.3 Outer Products

1. $\begin{bmatrix} 1 & 7 \\ 7 & 2 \end{bmatrix}$ (as long as the $(2,1)$ position is 7, the values on the diagonal entries could be any real number).

♠ 2. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 2 \\ 3 & 2 & 1 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 7 & 8 & 9 \\ 8 & 5 & 6 \\ 9 & 6 & 3 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 56 & 21 & 3 & 23 \\ 21 & 91 & 85 & 75 \\ 3 & 85 & 43 & 35 \\ 23 & 75 & 35 & 62 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 6 & 8 & 9 & 13 \\ 6 & 7 & 23 & 14 & 8 \\ 8 & 23 & 11 & 16 & 19 \\ 9 & 14 & 16 & 54 & 22 \\ 13 & 8 & 19 & 22 & 72 \end{bmatrix}$ (note that the $(1,1)$ position could be any real number)

6. $\begin{bmatrix} 1 & 7i & 5-2i \\ -7i & 3 & -9i \\ 5+2i & 9i & 2 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

8. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 3 & 5 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 0 & 1 & 2 \\ 3 & 4 & 5 & 6 \end{bmatrix}$

♠ 10. Hermitian

♠ 11. Hermitian

♠ 12. not Hermitian

♠ 13. Hermitian

14. $\begin{bmatrix} 0 & -2 \\ 2 & 0 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 3 & 9 & 3 \\ -3 & 0 & 5 & -2 \\ -9 & -5 & 0 & -9 \\ -3 & 2 & 9 & 0 \end{bmatrix}$

♠ 16. $\begin{bmatrix} -12 & 0 \\ 8 & 0 \end{bmatrix}$

♠ 17. $\begin{bmatrix} 8 & -2 & 4 \\ -8 & 2 & -4 \\ 4 & -1 & 2 \end{bmatrix}$

♠ 18. $\begin{bmatrix} 1 & -3 \\ 1 & -3 \\ 2 & -6 \end{bmatrix}$

19. $\boldsymbol{u} = \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} -4 \\ 0 \\ 1 \end{bmatrix}$

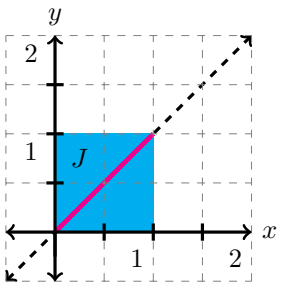
20. $\boldsymbol{u} = \begin{bmatrix} -2 \\ 1 \\ 4 \\ -1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} -7 \\ 2 \end{bmatrix}$

21. $\boldsymbol{u} = \begin{bmatrix} 3 \\ 3 \\ 6 \\ 9 \\ 4 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 3 \\ 2 \\ 4 \end{bmatrix}$

22. $\boldsymbol{u} = \begin{bmatrix} -9 \\ 0 \\ 0 \\ 8 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} -2 \\ 0 \\ 0 \\ -5 \end{bmatrix}$

♠ 23. idempotent ♠ 24. neither ♠ 25. nilpotent

♠ 26. The projection of the unit square is the line segment between the points $(0,0)$ and $(1,1)$. As such, it has no area because it is only a line segment. The notion of area is lost by projections.



27. (a) *Proof.* The result follows from the linearity of the transpose operator and $(A^\top)^\top = A$:

$$(A + A^\top)^\top = A^\top + (A^\top)^\top = A^\top + A = A + A^\top.$$ ∎

(b) *Proof.* The result follows from the linearity of the transpose operator and $(A^\top)^\top = A$:

$$(A - A^\top)^\top = A^\top + ((-A)^\top)^\top = A^\top + (-A) = -A + A^\top = -(A - A^\top).$$ ∎

(c) *Proof.* Let $S = \frac{1}{2}(A + A^\top)$ and $T = \frac{1}{2}(A - A^\top)$. Then

$$S + T = \frac{1}{2}(A + A^\top) + \frac{1}{2}(A - A^\top) = \frac{1}{2}(A + A) + \frac{1}{2}(A^\top - A^\top) = \frac{2}{2}A + 0 = A.$$

By part (a), we know that $A + A^\top$ is symmetric. Since the transpose operator is linear, any scalar multiple of a symmetric matrix is symmetric. Thus, S is symmetric. A similar argument using part (b) shows that T is skew-symmetric. Therefore, A can be decomposed into a sum of symmetric and skew-symmetric matrices. ∎

28. *Proof.* Let A be an idempotent matrix, that is, $A^2 = A$. Suppose that A is also nonsingular. Then

$$A^2 = A \quad \Rightarrow \quad A^{-1}(A^2) = A^{-1}(A) \quad \Rightarrow \quad I_n A = I_n \quad \Rightarrow \quad A = I_n.$$

Suppose next that A is nilpotent, that is, there exists some m such that $A^m = 0$. Let m be the smallest positive integer with this property. By the way of contradiction, suppose that A is also nonsingular. Then

$$A^m = 0 \quad \Rightarrow \quad A^{-1}(A^m) = A^{-1}(0) \quad \Rightarrow \quad I_n A^{m-1} = 0 \quad \Rightarrow \quad A^{m-1} = 0.$$

This contradicts the minimality of m. Therefore, A cannot be nilpotent and nonsingular. ∎

29. We say that a square matrix A is **skew-Hermitian** if $A^* = -A$. For example, the matrix $\begin{bmatrix} i & 3 & 4i \\ -3 & 0 & 1-2i \\ 4i & -1-2i & -7i \end{bmatrix}$ is skew-Hermitian.

4.4 Affine Transformations

♠ 1. $\cos^{-1}\left(\frac{2}{\sqrt{2}\sqrt{10}}\right) \approx 63.4°$ ♠ 2. $\cos^{-1}\left(\frac{-8}{\sqrt{10}\sqrt{14}}\right) \approx 132.5°$ ♠ 3. $\cos^{-1}\left(\frac{3}{\sqrt{30}}\right) \approx 56.8°$

♠ 4. rigid motion ♠ 5. rigid motion ♠ 6. not a rigid motion ♠ 7. not a rigid motion

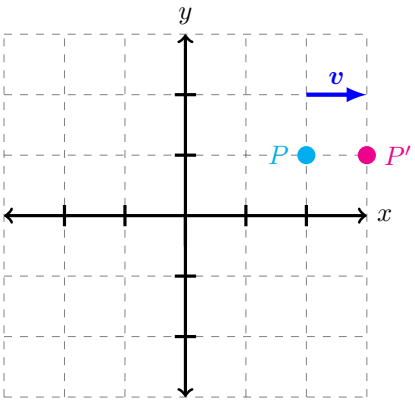
♠ 8. rigid motion

♠ 9. rigid motion

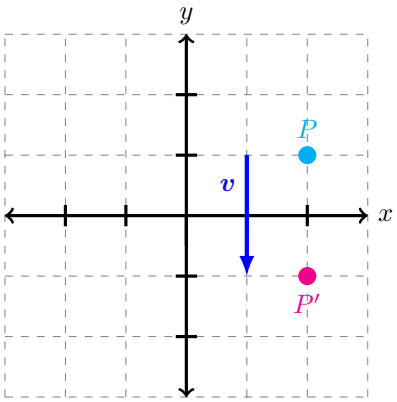
♠ 10. not a rigid motion

♠ 11. rigid motion

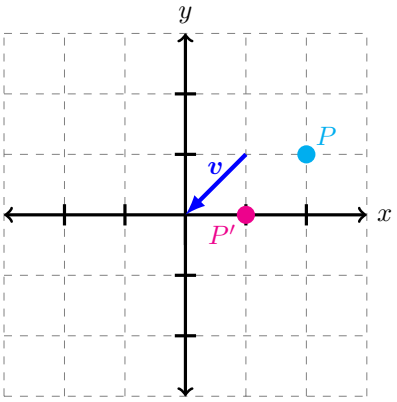
♠ 12.



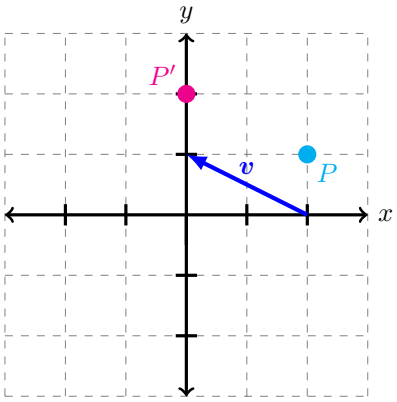
♠ 13.



♠ 14.



♠ 15.



16. $\begin{bmatrix} 4 \\ -5 \\ 8 \end{bmatrix}$

17. $\begin{bmatrix} 7 \\ 10 \\ -6 \end{bmatrix}$

18. Since $A \sim I_3$, the linear transformation associated to A is one-to-one and onto. Therefore, the affine transformation T is likewise one-to-one and onto.

♠ 19. $\begin{bmatrix} -4 \\ 20 \\ 14 \end{bmatrix}$

♠ 20. $\begin{bmatrix} -7 \\ 14 \\ -6 \end{bmatrix}$

♠ 21. Since $A \sim I_3$, the linear transformation associated to A is one-to-one and onto. Therefore, the affine transformation T is likewise one-to-one and onto.

22. $\begin{bmatrix} 23 \\ 10 \\ 33 \end{bmatrix}$

23. $\begin{bmatrix} -4 \\ 4 \\ 2 \end{bmatrix}$

24. Since $A \sim I_3$, the linear transformation associated to A is one-to-one and onto. Therefore, the affine transformation T is likewise one-to-one and onto.

4.5 Orthogonal Projections

♠ 1. $\boldsymbol{y} = -2\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} -2 \end{bmatrix}$

♠ 2. $\boldsymbol{y} = 5\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} - 6\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} 5 \\ -6 \end{bmatrix}$

♠ 3. $\boldsymbol{y} = \begin{bmatrix} 5 \\ 4 \\ -2 \end{bmatrix} + 6\begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$

♠ 4. $\boldsymbol{y} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 0 \end{bmatrix} - 2\begin{bmatrix} 3 \\ -1 \\ -5 \\ 1 \end{bmatrix} + 3\begin{bmatrix} -1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$

♠ 5. $\boldsymbol{y} = 2\begin{bmatrix} 1 \\ 6 \\ -6 \\ 0 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 7 \\ -2 \\ 0 \\ 5 \\ 5 \end{bmatrix}$, $[\boldsymbol{y}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$

♠ 6. $\text{proj}_W(\boldsymbol{y}) = \begin{bmatrix} 2/5 \\ 6/5 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} 2/5 \\ 6/5 \end{bmatrix} + \begin{bmatrix} 3/5 \\ -1/5 \end{bmatrix}$

♠ 7. $\text{proj}_W(\boldsymbol{y}) = \begin{bmatrix} 1/3 \\ 8/3 \\ 7/3 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} 1/3 \\ 8/3 \\ 7/3 \end{bmatrix} + \begin{bmatrix} 2/3 \\ -2/3 \\ 2/3 \end{bmatrix}$

♠ 8. $\text{proj}_W(\boldsymbol{y}) = \begin{bmatrix} 3 \\ -4/5 \\ 2/5 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} 3 \\ -4/5 \\ 2/5 \end{bmatrix} + \begin{bmatrix} 0 \\ 9/5 \\ 18/5 \end{bmatrix}$

♠ 9. $\text{proj}_W(\boldsymbol{y}) = \begin{bmatrix} 89/39 \\ 109/117 \\ 401/117 \\ 11/9 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} 89/39 \\ 109/117 \\ 401/117 \\ 11/9 \end{bmatrix} + \begin{bmatrix} -50/39 \\ 125/117 \\ -50/117 \\ 25/9 \end{bmatrix}$

♠ 10. $\text{proj}_W(\boldsymbol{y}) = \begin{bmatrix} 315/103 \\ -90/103 \\ 0 \\ 225/103 \\ 225/103 \end{bmatrix}$, $\boldsymbol{y} = \begin{bmatrix} 315/103 \\ -90/103 \\ 0 \\ 225/103 \\ 225/103 \end{bmatrix} + \begin{bmatrix} -212/103 \\ 193/103 \\ 2 \\ 84/103 \\ 290/103 \end{bmatrix}$

♠ 11. $\dfrac{\sqrt{10}}{5} \approx 0.632$

♠ 12. $\dfrac{2\sqrt{3}}{3} \approx 1.155$

♠ 13. $\dfrac{9\sqrt{5}}{5} \approx 4.025$

♠ 14. $\dfrac{25\sqrt{26}}{39} \approx 3.269$

♠ 15. $\dfrac{\sqrt{215785}}{103} \approx 4.510$

16. *Proof.*

$$\|\boldsymbol{u}+\boldsymbol{v}\|^2+\|\boldsymbol{u}-\boldsymbol{v}\|^2 = (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v})+(\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) = (\boldsymbol{u}\cdot\boldsymbol{u}+\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\boldsymbol{u}+\boldsymbol{v}\cdot\boldsymbol{v})+(\boldsymbol{u}\cdot\boldsymbol{u}-\boldsymbol{u}\cdot\boldsymbol{v}-\boldsymbol{v}\cdot\boldsymbol{u}+\boldsymbol{v}\cdot\boldsymbol{v})$$
$$= 2(\boldsymbol{u}\cdot\boldsymbol{u}) + 2(\boldsymbol{v}\cdot\boldsymbol{v}) = 2\|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|^2. \qquad \blacksquare$$

4.6 The Fundamental Theorem of Linear Algebra

1. 2×5 2. 4×4 3. 3×3 4. 4×2

♠ 5. $\left\{ \begin{bmatrix} 7 \\ 1 \\ -2 \end{bmatrix} \right\}$ ♠ 6. $\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ -4 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ ♠ 7. $\left\{ \begin{bmatrix} -71 \\ 79 \\ 70 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 5 \end{bmatrix} \right\}$

♠ 8. $\begin{bmatrix} 4/13 \\ 6/13 \end{bmatrix}$ ♠ 9. $\begin{bmatrix} -8/7, \\ 4/7 \\ 16/7 \end{bmatrix}$ ♠ 10. $\begin{bmatrix} 1/5 \\ 1/5 \\ -7/5 \\ 8/5 \end{bmatrix}$

4.7 The Gram-Schmidt Algorithm

♠ 1. $\left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}$ ♠ 2. $\left\{ \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}, \begin{bmatrix} 2/3 \\ 1/3 \\ -2/3 \end{bmatrix} \right\}$ ♠ 3. $\left\{ \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}, \begin{bmatrix} -4/5 \\ 3/5 \end{bmatrix} \right\}$

♠ 4. $\left\{ \begin{bmatrix} 1/\sqrt{11} \\ 0 \\ 3/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \begin{bmatrix} 6/\sqrt{209} \\ 11/\sqrt{209} \\ -4/\sqrt{209} \\ 6/\sqrt{209} \end{bmatrix} \right\}$ ♠ 5. $\left\{ \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}, \begin{bmatrix} 11/\sqrt{266} \\ 8/\sqrt{266} \\ -9/\sqrt{266} \end{bmatrix}, \begin{bmatrix} 3/\sqrt{19} \\ -3/\sqrt{19} \\ 1/\sqrt{19} \end{bmatrix} \right\}$

♠ 6. $\left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -i \\ 2i \end{bmatrix}, \frac{1}{\sqrt{78}} \begin{bmatrix} 2+i \\ 7-2i \\ 4-2i \end{bmatrix} \right\}$ ♠ 7. $\left\{ \frac{1}{2} \begin{bmatrix} 1 \\ i \\ -1 \\ -i \end{bmatrix}, \frac{1}{4\sqrt{7}} \begin{bmatrix} 3-i \\ 5+i \\ 5+i \\ 7-i \end{bmatrix}, \frac{1}{2\sqrt{1302}} \begin{bmatrix} 31-11i \\ -43+i \\ 41+15i \\ -17-9i \end{bmatrix} \right\}$

♠ 8. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2/\sqrt{2} & 1/\sqrt{2} \\ 0 & 3/\sqrt{2} \end{bmatrix}$ ♠ 9. $\begin{bmatrix} 2/3 & 2/3 \\ -2/3 & 1/3 \\ 1/3 & -2/3 \end{bmatrix} \begin{bmatrix} 6 & -3 \\ 0 & 6 \end{bmatrix}$

♠ 10. $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2/\sqrt{2} & 3/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{bmatrix}$

♠ 11. $\begin{bmatrix} 1/\sqrt{11} & 6/\sqrt{209} \\ 0 & 11/\sqrt{209} \\ 3/\sqrt{11} & -4/\sqrt{209} \\ 1/\sqrt{11} & 6/\sqrt{209} \end{bmatrix} \begin{bmatrix} \sqrt{11} & \frac{16}{\sqrt{11}} \\ 0 & \frac{19}{\sqrt{209}} \end{bmatrix}$

♠ 12. $\begin{bmatrix} 1/\sqrt{14} & 11/\sqrt{266} & 3/\sqrt{19} \\ 2/\sqrt{14} & 8/\sqrt{266} & -3/\sqrt{19} \\ 3/\sqrt{14} & -9/\sqrt{266} & 1/\sqrt{19} \end{bmatrix} \begin{bmatrix} 14/\sqrt{14} & -5/\sqrt{14} & 18/\sqrt{14} \\ 0 & 57/\sqrt{266} & 30/\sqrt{266} \\ 0 & 0 & 3/\sqrt{19} \end{bmatrix}$

4.8 The Least-Squares Problem

♠ 1. $\frac{1}{11}\begin{bmatrix} 20 \\ -8 \end{bmatrix}$

♠ 2. $\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 14 \\ 30 \\ 26 \end{bmatrix}$

♠ 4. $A^\top \boldsymbol{b} = \begin{bmatrix} 1 \\ 10 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 14 & -3 \\ -3 & 21 \end{bmatrix}$, $\widehat{\boldsymbol{x}} = \frac{1}{285}\begin{bmatrix} 51 \\ 143 \end{bmatrix}$, $\widehat{\boldsymbol{b}} = \frac{1}{285}\begin{bmatrix} -92 \\ 439 \\ 470 \end{bmatrix}$

♠ 5. $A^\top \boldsymbol{b} = \begin{bmatrix} -6 \\ -4 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 42 & 0 \\ 0 & 14 \end{bmatrix}$, $\widehat{\boldsymbol{x}} = \frac{1}{7}\begin{bmatrix} -1 \\ -2 \end{bmatrix}$, $\widehat{\boldsymbol{b}} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$

♠ 6. $A = \begin{bmatrix} -1 & 2 \\ 2 & 2 \\ 1 & 4 \end{bmatrix}$, $A^\top \boldsymbol{b} = \begin{bmatrix} -12 \\ -6 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 6 & 6 \\ 6 & 24 \end{bmatrix}$, $\widehat{\boldsymbol{x}} = \frac{1}{3}\begin{bmatrix} -7 \\ 1 \end{bmatrix}$, $\widehat{\boldsymbol{b}} = \begin{bmatrix} 3 \\ -4 \\ -1 \end{bmatrix}$

♠ 7. $A = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \end{bmatrix}$, $A^\top \boldsymbol{b} = \begin{bmatrix} 30 \\ 21 \\ -21 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 7 & 4 & -6 \\ 4 & 3 & -3 \\ -6 & -3 & 6 \end{bmatrix}$, $\widehat{\boldsymbol{x}} = \begin{bmatrix} 6 \\ 3 \\ 4 \end{bmatrix}$, $\widehat{\boldsymbol{b}} = \begin{bmatrix} 7 \\ 2 \\ 9 \\ 5 \end{bmatrix}$

♠ 8. $W = \text{Span}\left\{\begin{bmatrix} 1 \\ 0 \\ -5 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix}\right\}$, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -5 & 3 \end{bmatrix}$, $A(A^\top A)^{-1}A^\top = A\begin{bmatrix} 26 & -15 \\ -15 & 10 \end{bmatrix}^{-1} A^\top =$

$$\frac{1}{35}A\begin{bmatrix}10 & 15\\ 15 & 26\end{bmatrix}A^{\top}=\frac{1}{35}A\begin{bmatrix}10 & 15 & -5\\ 15 & 26 & 3\end{bmatrix}=\frac{1}{35}\begin{bmatrix}10 & 15 & -5\\ 15 & 26 & 3\\ -5 & 3 & 34\end{bmatrix}$$

♠ 9. $W=\text{Span}\left\{\begin{bmatrix}2\\ -1\\ 4\end{bmatrix}\right\}$, $A=\begin{bmatrix}2\\ -1\\ 4\end{bmatrix}$ $A(A^{\top}A)^{-1}A^{\top}=A\begin{bmatrix}21\end{bmatrix}^{-1}A^{\top}=\frac{1}{21}AA^{\top}=\frac{1}{21}\begin{bmatrix}4 & -2 & 8\\ -2 & 1 & -4\\ 8 & -4 & 16\end{bmatrix}$

10. Hint: By Theorem 4.8.5, $\widehat{\boldsymbol{x}}$ is the linear solution to the matrix equation $A^{\top}A\boldsymbol{x}=A^{\top}\boldsymbol{b}$. Since A is invertible, show that the solution set of $A\boldsymbol{x}=\boldsymbol{b}$ and $A^{\top}A\boldsymbol{x}=A^{\top}\boldsymbol{b}$ are the same.

11. Hint: You may use the fact $Q^{\top}Q=I_n$. Then substitute $A=QR$ into the equation from Theorem 4.8.5. Also, although we know that R is nonsingular and we can use R^{-1}, we do not know that Q is nonsingular. Thus, we cannot use Q^{-1} because it might not be defined.

Chapter 5 : Determinants

5.1 Introduction to Determinants

1. $\begin{bmatrix}5 & 12\\ 9 & -6\end{bmatrix}$ 2. $\begin{bmatrix}5 & 7\\ 0 & 3\end{bmatrix}$ 3. $\begin{bmatrix}3 & 4\\ -1 & -6\end{bmatrix}$ 4. $\begin{bmatrix}5 & 7\\ 9 & -1\end{bmatrix}$ 5. $\begin{bmatrix}0 & 4\\ 9 & -6\end{bmatrix}$

♠ 6. $\begin{bmatrix}-2 & -1 & 3\\ 1 & 3 & 6\\ -4 & 8 & 0\end{bmatrix}$ ♠ 7. $\begin{bmatrix}0 & -2 & 3\\ 1 & 1 & 6\\ 7 & -4 & 0\end{bmatrix}$ ♠ 8. $\begin{bmatrix}2 & 3 & 4\\ -2 & -1 & 3\\ -4 & 8 & 0\end{bmatrix}$ ♠ 9. $\begin{bmatrix}1 & 2 & 4\\ 0 & -2 & 3\\ 7 & -4 & 0\end{bmatrix}$

♠ 10. $\begin{bmatrix}1 & 2 & 3\\ 0 & -2 & -1\\ 7 & -4 & 8\end{bmatrix}$ 11. $\begin{bmatrix}1 & 3 & 4\\ -2 & 0 & -4\\ -1 & 2 & 5\end{bmatrix}$ 12. $\begin{bmatrix}1 & 2 & 4\\ -2 & 4 & -4\\ -1 & 3 & 5\end{bmatrix}$

13. $\begin{bmatrix}1 & 3 & 4\\ 0 & 3 & -1\\ -2 & 0 & -4\end{bmatrix}$ 14. $\begin{bmatrix}i & 3i-2 & 2+5i\\ 2+3i & 5i-1 & 1+9i\\ 7+2i & i-3 & 2+4i\end{bmatrix}$ 15. $\begin{bmatrix}i & 3i-2 & 7-i\\ 3+i & 4 & 9-3i\\ 7+2i & i-3 & 1+2i\end{bmatrix}$

16. $\begin{bmatrix}3+i & 4 & 7i-5\\ 2+3i & 5i-1 & 1+9i\\ 7+2i & i-3 & 2+4i\end{bmatrix}$

♠ 17. -2

♠ 18. $1+2i$

♠ 19. 6

♠ 20. 9

♠ 21. -16

♠ 22. 3

♠ 23. 4

♠ 24. 0 25. 12 26. $\pi^5 - 25e^2$ ♠ 27. 4 ♠ 28. 70 ♠ 29. 42

5.2 Properties of Determinants

1. -5 ♠ 2. -2 ♠ 3. 10 4. -10 5. 7

♠ 6. 6 ♠ 7. 52 8. -54 9. 52 10. $-9+i$

♠ 11. 4 12. 24 13. 0 14. 126 ♠ 15. 1 16. 27

♠ 17. 6 ♠ 18. 127 19. 729 ♠ 20. $\text{rank}(A) = 5$, $\text{nullity}(A) = 0$ ♠ 21. $\text{corank}(A) = 5$, $\text{conullity}(A) = 0$

♠ 22. Yes, it is consistent for any $\boldsymbol{b}$. $A\boldsymbol{x} = \boldsymbol{b}$ has no free variables. $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution.

♠ 23. I_5

♠ 24. $\ker(T) = \{\mathbf{0}\}$, $\text{im}(T) = \mathbb{R}^5$. Yes, T is both one-to-one and onto.

25. *Proof.* Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$. Then $\det(A) = ad - bc$ and $\det(B) = eh - fg$. Hence,

$$\det(A)\det(B) = (ad-bc)(eh-fg) = adeh - adfg - bceh + bcfg.$$

On the other hand, $AB = \begin{bmatrix} ae+bf & af+bh \\ ce+dg & cf+dh \end{bmatrix}$. Hence,

$$\begin{aligned}\det(AB) = (ae+bf)(cf+dh)-(af+bh)(ce+dg) &= aecf+aedh+bfch+bfdh-afce-afdg-bhce-bhdg \\ &= adeh - adfg - bceh + bcfg = \det(A)\det(B). \quad \blacksquare\end{aligned}$$

5.3 Cramer's Rule

1. $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ ♠ 2. $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ ♠ 3. $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ ♠ 4. $\begin{bmatrix} 12/17 \\ -1/17 \\ 5/17 \end{bmatrix}$ ♠ 5. $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}$

♠ 6. $\text{adj}(A) = \begin{bmatrix} -3 & -3 \\ -2 & 1 \end{bmatrix}$,

$A(\text{adj}\, A) = \begin{bmatrix} -9 & 0 \\ 0 & -9 \end{bmatrix}$

♠ 7. $\text{adj}(A) = \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}$,

$A(\text{adj}\, A) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

♠ 8. $\text{adj}(A) = \begin{bmatrix} -6 & 10 & 19 \\ 9 & -15 & -3 \\ 6 & 7 & -2 \end{bmatrix}$,

$A(\text{adj}\, A) = \begin{bmatrix} 51 & 0 & 0 \\ 0 & 51 & 0 \\ 0 & 0 & 51 \end{bmatrix}$

♠ 9. $\operatorname{adj}(A) = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 2 & 1 & 2 & 1 \end{bmatrix}$, $A(\operatorname{adj} A) = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

10. $\det(A) = 1$, $A^{-1} = \begin{bmatrix} 4 & 13 & 5 \\ -3 & -11 & -4 \\ 1 & 3 & 1 \end{bmatrix}$

11. $\det(A) = 4$, $A^{-1} = \begin{bmatrix} 2 & 3/2 & 2 \\ 0 & -1/2 & 3/2 \\ 1 & 1 & 0 \end{bmatrix}$

12. $\det(A) = -158$, $A^{-1} = \frac{1}{158}\begin{bmatrix} 10 & -30 & -2 \\ -11 & -46 & 18 \\ 19 & 22 & 12 \end{bmatrix}$

5.4 Cross Products

♠ 1. $\begin{bmatrix} -3 \\ 7 \\ 5 \end{bmatrix}$ ♠ 2. $\begin{bmatrix} -26 \\ 9 \\ 2 \end{bmatrix}$ ♠ 3. $\begin{bmatrix} 4 \\ -6 \\ 2 \end{bmatrix}$ ♠ 4. $\begin{bmatrix} 1 \\ -5 \\ 1 \end{bmatrix}$ ♠ 5. $\begin{bmatrix} 19 \\ -34 \\ 12 \end{bmatrix}$

♠ 6. $x + y + 6z = 5$ ♠ 7. $x + 2y + 3z + 4w = 5$ ♠ 8. $\begin{cases} -x + y = 1 \\ 3x - 3z + 4w = 10 \end{cases}$

9. *Proof.* To prove part (ix),

$$\boldsymbol{u}\times(\boldsymbol{v}+\boldsymbol{w}) = \boldsymbol{u}\times\begin{bmatrix} v_1+w_1 \\ v_2+w_2 \\ v_3+w_3 \end{bmatrix} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ u_1 & u_2 & u_3 \\ v_1+w_1 & v_2+w_2 & v_3+w_3 \end{vmatrix} = \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} + \begin{vmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \\ u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$
$$= (\boldsymbol{u}\times\boldsymbol{v}) + (\boldsymbol{u}\times\boldsymbol{w}),$$

where the second to last equality follows from multilinearity of the determinant. ∎

10. *Proof.* To prove part (xii),

$$\boldsymbol{u}\times\boldsymbol{u} = \left(\begin{vmatrix} u_2 & u_3 \\ u_2 & u_3 \end{vmatrix}, -\begin{vmatrix} u_1 & u_3 \\ u_1 & u_3 \end{vmatrix}, \begin{vmatrix} u_1 & u_2 \\ u_1 & u_2 \end{vmatrix} \right) = (0, -0, 0) = \boldsymbol{0}.$$ ∎

Chapter 6 : Eigenvalues

6.1 Eigenvalues and Eigenvectors

1. yes, $\lambda = 5$ 2. no 3. no ♠ 4. yes, $\lambda = 2$ ♠ 5. no ♠ 6. yes, $\lambda = -1$

♠ 7. yes, $\lambda = 1$ ♠ 8. yes, $\lambda = 0$ ♠ 9. no ♠ 10. $c\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ♠ 11. $c\begin{bmatrix} 3 \\ 1 \end{bmatrix}$ ♠ 12. $c\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

♠ 13. yes ♠ 14. yes ♠ 15. no

♠ 16. $\lambda = 5$, mult $= 1$, $\left\{\begin{bmatrix} 2 \\ -4 \\ 1 \end{bmatrix}\right\}$; $\lambda = -3$, mult $= 1$, $\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}\right\}$

♠ 17. $\lambda = -3$, mult $= 1$, $\left\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\}$; $\lambda = 2$, mult $= 2$, $\left\{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}\right\}$

18. $\lambda = 1$, mult $= 1$, $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right\}$; $\lambda = 2$, mult $= 1$, $\left\{\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right\}$

19. Hint: If $A\boldsymbol{x} = \lambda\boldsymbol{x}$, then what can we say about $A^2\boldsymbol{x} = A(A\boldsymbol{x})$? Can we generalize this observation?

20. Hint: If A is idempotent, then $A^2 = A$. Use Exercise 19.

21. Hint: If A is nilpotent, then $A^n = 0$ for some n. Use Exercise 19.

6.2 The Characteristic Polynomial

♠ 1. $\lambda = 4$, $\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$; $\lambda = -2$, $\left\{\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right\}$ ♠ 2. $\lambda = 5$, $\left\{\begin{bmatrix} -1 \\ 1 \end{bmatrix}\right\}$ ♠ 3. $\lambda = 3$, $\left\{\begin{bmatrix} -3 \\ 2 \end{bmatrix}\right\}$; $\lambda = -1$, $\left\{\begin{bmatrix} -4 \\ 3 \end{bmatrix}\right\}$

♠ 4. $\lambda = 2+i$, $\left\{\begin{bmatrix} 2+i \\ 1 \end{bmatrix}\right\}$; $\lambda = 2-i$, $\left\{\begin{bmatrix} 2-i \\ 1 \end{bmatrix}\right\}$ ♠ 5. $\lambda = 3+2i$, $\left\{\begin{bmatrix} -2+i \\ 2 \end{bmatrix}\right\}$; $\lambda = 3-2i$, $\left\{\begin{bmatrix} -2-i \\ 2 \end{bmatrix}\right\}$ ♠ 6. $\lambda = 4+i$, $\left\{\begin{bmatrix} 1+i \\ 1 \end{bmatrix}\right\}$; $\lambda = 4-i$, $\left\{\begin{bmatrix} 1-i \\ 1 \end{bmatrix}\right\}$

♠ 7. $\lambda = 1$, $\left\{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right\}$; $\lambda = 0$, $\left\{\begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\}$; $\lambda = 2$, $\left\{\begin{bmatrix} 1 \\ -3 \\ -1 \end{bmatrix}\right\}$

♠ 8. $\lambda = 2$, $\left\{\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right\}$; $\lambda = -3$, $\left\{\begin{bmatrix} 1 \\ -3 \\ -1 \end{bmatrix}\right\}$

♠ 9. $\lambda = 0, \left\{ \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$

♠ 10. A and B are not similar since $\text{tr}(A) = 5 \neq 3 = \text{tr}(B)$

♠ 11. A and B are not similar since $\det(A) = 4 \neq 6 = \det(B)$

6.3 Diagonalization

♠ 1. $\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -2 & 1 \end{bmatrix}$

♠ 3. $\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 \\ -1 & 3 & -1 \\ 1 & -2 & 1 \end{bmatrix}$

♠ 4. $\begin{bmatrix} 2 & 0 & 2 & 5 \\ 1 & 0 & 1 & 3 \\ 4 & 1 & 0 & 12 \\ 5 & 0 & 6 & 15 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 & -1 \\ 0 & -24 & 1 & 4 \\ 0 & -5 & 0 & 1 \\ -1 & 2 & 0 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 4 & 2 & -1 \\ 1 & -3 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 18 & 0 & 0 \\ 0 & -17 & 0 \\ 0 & 0 & -2 \end{bmatrix} \left(\frac{1}{21} \begin{bmatrix} 3 & 3 & 3 \\ 1 & -6 & 1 \\ -7 & 0 & 14 \end{bmatrix} \right)$

6.4 Orthogonal Diagonalization

♠ 1. $\begin{bmatrix} 5/\sqrt{26} & -1/\sqrt{26} \\ 1/\sqrt{26} & 5/\sqrt{26} \end{bmatrix} \begin{bmatrix} 18 & 0 \\ 0 & -8 \end{bmatrix} \begin{bmatrix} 5/\sqrt{26} & 1/\sqrt{26} \\ -1/\sqrt{26} & 5/\sqrt{26} \end{bmatrix}$

♠ 2. $\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 30 & 0 & 0 \\ 0 & -12 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{6} & 1/\sqrt{6} & -2/\sqrt{6} \end{bmatrix}$

♠ 3. $\begin{bmatrix} 1/2 & 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{12} \\ -1/2 & 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{12} \\ 1/2 & 0 & 2/\sqrt{6} & 1/\sqrt{12} \\ -1/2 & 0 & 0 & 3/\sqrt{12} \end{bmatrix} \begin{bmatrix} -4 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} 1/2 & -1/2 & 1/2 & -1/2 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} & 0 \\ 1/\sqrt{12} & -1/\sqrt{12} & 1/\sqrt{12} & 3/\sqrt{12} \end{bmatrix}$

♠ 4. $\begin{bmatrix} (-1+i)/\sqrt{3} & (1-i)/\sqrt{6} \\ 1/\sqrt{3} & 2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} (-1-i)/\sqrt{3} & 1/\sqrt{3} \\ (1+i)/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}$

♠ 5. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & (1-i)/\sqrt{3} & (-1+i)/\sqrt{6} \\ 0 & 1/\sqrt{3} & 2/\sqrt{6} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & (1+i)/\sqrt{3} & 1/\sqrt{3} \\ 0 & (-1-i)/\sqrt{6} & 2/\sqrt{6} \end{bmatrix}$

♠ 6. $18 \begin{bmatrix} 25/26 & 5/26 \\ 5/26 & 1/26 \end{bmatrix} - 8 \begin{bmatrix} 1/26 & -5/26 \\ -5/26 & 25/26 \end{bmatrix}$

♠ 7. $30 \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix} - 12 \begin{bmatrix} 1/2 & -1/2 & 0 \\ -1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{bmatrix} + 6 \begin{bmatrix} 1/6 & 1/6 & -1/3 \\ 1/6 & 1/6 & -1/3 \\ -1/3 & -1/3 & 2/3 \end{bmatrix}$

♠ 8. $$-4\begin{bmatrix} 1/4 & -1/4 & 1/4 & -1/4 \\ -1/4 & 1/4 & -1/4 & 1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ -1/4 & 1/4 & -1/4 & 1/4 \end{bmatrix} + 8\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 8\begin{bmatrix} 1/6 & -1/6 & -1/3 & 0 \\ -1/6 & 1/6 & 1/3 & 0 \\ -1/3 & 1/3 & 2/3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + 8\begin{bmatrix} 1/12 & -1/12 & 1/12 & 1/4 \\ -1/12 & 1/12 & -1/12 & -1/4 \\ 1/12 & -1/12 & 1/12 & 1/4 \\ 1/4 & -1/4 & 1/4 & 3/4 \end{bmatrix}$$
$$= -4\begin{bmatrix} 1/4 & -1/4 & 1/4 & -1/4 \\ -1/4 & 1/4 & -1/4 & 1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ -1/4 & 1/4 & -1/4 & 1/4 \end{bmatrix} + 8\begin{bmatrix} 3/4 & 1/4 & -1/4 & 1/4 \\ 1/4 & 3/4 & 1/4 & -1/4 \\ -1/4 & 1/4 & 3/4 & 1/4 \\ 1/4 & -1/4 & 1/4 & 3/4 \end{bmatrix}$$

9. Hint: Recall that $(A^{-1})^\top = (A^\top)^{-1}$.

10. Hint: Recall that $(AB)^\top = B^\top A^\top$. This is the shoes-socks property. Doesn't another matrix operation also have the shoes-socks property?

11. Hint: Recall that $\det(A^\top) = \det(A)$.

12. Hint: You may use the fact that $(U\boldsymbol{x}) \cdot (U\boldsymbol{y}) = \boldsymbol{x} \cdot \boldsymbol{y}$ for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ if and only if U is orthogonal.

6.5 Similarity and Linear Transformations

♠ 1. $$\begin{bmatrix} -7 & -1 \\ -5 & -1 \\ -11 & -1 \end{bmatrix}$$

♠ 2. $$\begin{bmatrix} 0 & 8 & -5 \\ 0 & 11 & -7 \end{bmatrix}$$

♠ 3. $$\begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$$